The World As I Found It. A Subjectivist Metaphysics of the Mental

Gionvanni Merlo

THE WORLD AS I FOUND IT

A Subjectivist Metaphysics of the Mental

Giovanni Merlo

Supervisors:

Prof Sven Rosenkranz

Prof John Hawthorne

Tutor:

Prof Sven Rosenkranz

PhD in Analytic Philosophy (HDC04)

Department of Logic, History and the Philosophy of Science

University of Barcelona

2015

For my parents

R E S U M E N

La primera parte de esta tesis articula y defiende la Teoría Subjetivista de lo Mental. Según esta teoría, mis estados mentales son esencialmente diferentes de los estados mentales de los demás, pero el hecho de que lo son no es un hecho objetivo, sino que un hecho subjetivo. El Capítulo 1 explica qué significa que un hecho sea subjetivo, qué tipo de diferencia hay entre mis estados mentales y los estados mentales de los demás y qué tipo de intuiciones me llevan a creer que haya esta diferencia. El Capítulo 2 defiende la Teoría Subjetivista de lo Mental frente a objeciones y discute algunas de sus implicaciones más importantes. En la segunda parte de la tesis, voy a examinar las ventajas que la Teoría Subjetivista de lo Mental ofrece cuando se trata de dar cuenta de tres fenómenos característicos de nuestro conocimiento de los hechos mentales. El Capítulo 3 trata de la asimetría entre el conocimiento de nuestra propia mente y el conocimiento de otras mentes. Se argumenta que, si la Teoría Subjetivista de lo Mental es correcta, es posible explicar esta asimetría sin asumir que el conocimiento de nuestra propia mente y el conocimiento de otras mentes están basados en diferentes formas de conocimiento. El Capítulo 4 se centra en la "inmunidad al error mediante identificación equivocada" de las auto-atribuciones mentales. Lo que voy a mostrar es que, aunque haya habido una tendencia a explicar este fenómeno diciendo (como Lichtenberg y Wittgenstein) que el contenido de la auto-atribuciones mentales es general y no particular, esta estrategia no puede funcionar a menos que la Teoría Subjetivista de lo Mental sea correcta. El Capítulo 5 se ocupa del conocimiento experiencial. Se argumenta que, si las verdades experienciales son consideradas como verdades objetivas, es difícil explicar porqué el conocimiento experiencial tiene que ser conocimiento de 'primera mano'. El problema desaparece si las verdades experienciales son consideradas como verdades subjetivas.

A B S T R A C T

The first part of this thesis articulates and defends the Subjectivist View of the Mental. According to this view, my mental states are essentially different from the mental states of everyone else, but the fact that they are is a subjective fact, rather than an objective one. Chapter 1 explains what it takes for a fact to be subjective, what kind of difference holds between my mental states and everyone else's mental states and what kind of intuitions lead me to believe that there is such a difference. Chapter 2 defends the Subjectivist View of the Mental from objections and discusses some of its most significant implications. In the second part of the thesis, I explore the advantages that the Subjectivist View of the Mental offers when it comes to accounting for three basic features of our knowledge of the mental. Chapter 3 deals with the asymmetry between self-knowledge and knowledge of other minds. It is argued that, if the Subjectivist View of the Mental is true, we can explain why this asymmetry holds without assuming that self-knowledge and knowledge of other minds are involved with substantially different ways of knowing. Chapter 4 focuses on the immunity to error through misidentification of mental self-ascriptions. I show that, while there has been a temptation to explain this phenomenon by following Lichtenberg and Wittgenstein in taking the content of mental self-ascriptions to be general rather than particular, this proposal cannot be made to work unless something like the Subjectivist View of the Mental is true. Chapter 5 is concerned with experiential knowledge. I argue that, if experiential truths are taken to be objective truths, it is not easy to see why experiential knowledge should have the peculiar 'first-hand' character it has. The problem disappears if experiential truths are regarded as subjective rather than objective.

C O N T E N T S

Acknowledgements

I would like to thank all of those who have discussed with me the ideas in this dissertation, or provided me with comments on previous drafts of its chapters. In particular, thanks to Tom Avery, Derek Ball, Ned Block, Philipp Blum, Emiliano Boccardi, David Chalmers, Sam Coleman, Fabrice Correia, Alessandro Di Nicola, Katalin Farkas, Filippo Ferrari, Patrick Greenough, Katherine Hawley, Kit Fine, Martin Lipman, Guy Longworth, Aidan McGlynn, Daniel Morgan, Kevin Mulligan, Michele Ombrato, Peter Pagin, Giulia Pravato, Simon Prosser, François Recanati, Howard Robinson, Gonzalo Rodriguez-Pereyra, Susanna Schellenberg, Moritz Schulz, Barry Smith, Giuliano Torrengo, Peter Unger, Assaf Weksler, Richard Woodward and Crispin Wright.

I am very grateful to all the members of *LOGOS* and the philosophical community at the University of Barcelona – thanks for many stimulating discussions and for providing me with an exceptional research environment. Thanks also to the members of the other research groups who hosted me during my academic visits abroad: *Eidos* at the University of Geneva, the *Northern Institute of Philosophy* at the University of Aberdeen, *Arché* at the University of St Andrews and the Philosophy Department at New York University.

Above all, I would like to thank my supervisor, Sven Rosenkranz, who has helped me with incredible patience, generosity and dedication. Without his insightful comments and challenging criticisms, this thesis would have been so much worse than it is. I owe a very special thank to my co-supervisor, John Hawthorne, for his invaluable advice and encouragement, and for countless illuminating conversations over the years.

I feel privileged to have been a pre-doctoral fellow with *PETAF*, the Marie Curie Initial Training Network on Perspectival Thoughts and Facts, which was funded by the European Community's Seventh Framework Programme FP7/2007–2013 under grant agreement no. FP7-238128. Part of the research leading to this dissertation was carried out while I was holding a visiting fellowship at the New York University Center for Mind, Brain and Consciousness. Thanks to Ned Block and David Chalmers for making my visit possible.

PREFACE

Every time I think about myself and my place in reality, I feel torn between two opposing impulses. A part of me wants to say that there is something special about the individual I call "myself" – something that gives that individual the protagonist role on the stage of the world, while relegating to a secondary position all the individuals I call "others". But the rest of me knows that the individuals I call "others" are perfectly entitled to regard me the way I regard them – whatever privilege I claim for myself and deny to them, they can, with equal right, claim for themselves and deny to me. Can I really take myself to be in any way special, then?

The view I articulate in this thesis aims to offer a way out of this predicament. It says that, due to the very nature of my mental states, I am, indeed, special in comparison with other subjects. But it also says that I should regard the fact of my specialness as a subjective fact – a fact that obtains from my point of view, but does not obtain from the point of view of other subjects. Because it makes essential use of the notion of a subjective fact I call this view a *subjectivist* view. And because it locates the source of my specialness in the nature of my mental states I call it a subjectivist view *of the mental*.

The first chapter of this thesis will be devoted to explaining what I mean by "subjective fact" and what kind of intuitions led me to embrace the Subjectivist View of the Mental. In Chapter 2, I will defend the view in the face of some objections – from the charge of being a form of solipsism in disguise to the accusation of creating seemingly insuperable disagreement among its supporters.

In the second part of the thesis, the Subjectivist View of the Mental will be put to work. My focus will be on three peculiar features of our knowledge of the mental – the asymmetry between self-knowledge and knowledge of other minds (Chapter 3), the immunity to error through misidentification of the judgments we form about our mental states (Chapter 4) and the essentially 'firsthand' character of experiential knowledge (Chapter 5). My claim will be that we cannot easily explain why our knowledge of the mental exhibits these features unless something like the Subjectivist View of the Mental is true.

In a way, then, this thesis offers the Subjectivist View of the Mental not only as a way of doing justice to some deep-rooted intuitions, but also as the best response to what Christopher Peacocke would call an 'Integration Challenge' – the challenge of "providing, for a given area, a simultaneously acceptable metaphysics and epistemology" (1999, 1) or, equivalently, of "[reconciling] a plausible account of what is involved in the truth of statements of a given kind with a credible account of how we can know those statements, when we do know them" (ibid.). The particular 'area' I will be interested in is the domain of mental facts. The particular challenge I will focus on is that of reconciling a plausible metaphysics of the mental – i.e. a plausible account of the nature of mental facts – with a credible epistemology of mental truths – i.e. a credible account of how we know mental facts, when we know them.

PART I

CHAPTER 1:

Subjectivism and the Mental

There does not seem to be much point in the notion of a subjective fact – something that is a *fact* without being *objectively* the case. Among the things we call "facts" it is hard to find anything worth describing as "subjective". For example, take the fact that I am sitting right now. While this fact concerns me as opposed to anyone else, no one would describe it as a "subjective" fact: that I am sitting is true from my point of view just as much as it is true from your or anyone else's point of view – in a word, it is objectively true. Conversely, among the things we describe as "subjective" it is hard to find anything worth calling a "fact". For example, some have suggested that whether chocolate is tasty should be regarded as a subjective matter: that chocolate is tasty is true by the standards of some people and false by those of others. But even those who find chocolate tasty would feel nervous saying that it's a "fact" that chocolate is tasty. After all, there seems to be no absolute fact of the matter as to whether chocolate is tasty; and if there is no *absolute* fact of the matter as to whether chocolate is tasty, how could there be such a thing as the *fact* that chocolate is tasty?

In effect, it may be suggested that the notion of a subjective fact is not just pointless, but also incoherent. For when we describe something as "subjective" we mean that it reflects a particular point of view on reality. But when we describe something as a "fact", we mean that it reflects the way reality is in and of itself. So how could something be a fact and, at the same time, be subjective? The very idea of a subjective fact seems to be a contradiction in

terms.[1]

Given these premises, it is no wonder that *Subjectivism* – the view that reality is only subjectively the way it is – has received little attention and even less support from contemporary metaphysicians: if there are no subjective facts, every question concerning how reality is in and of itself must have an objective answer. And this means that, whichever way reality is, it must be objectively that way.

In rejecting both the premises and the conclusions of the foregoing line of reasoning, this introductory chapter sets the agenda for the rest of my dissertation. I will begin by arguing that – given certain well-known analogies between subjectivity, modality and time – there is no good reason to regard the notion of a subjective fact as contradictory or incoherent (§1). I will then present my grounds for thinking that there are, indeed, subjective facts (§ 2). On the version of Subjectivism I will put forward, subjective facts do not concern just *me*: they concern every individual endowed with a mental life. And they do not (at least, directly) concern evaluative properties like the property of being tasty: they concern the distribution, among the individuals there are, of certain mental properties. I will therefore call my preferred version of Subjectivism the "Subjectivist View of the Mental". § 3 offers a precise statement of the view. § 4 discusses some of its historical predecessors and compares it with other versions of Subjectivism.

1 This is what, among others, Moore (1997, 45-50) argues.

1. Subjectivism

I call something a "subjective fact" when it is a fact, but it is not objectively the case. Strange as it may appear at first, this notion is not without parallels in other areas of metaphysics. The modal analogue of a subjective fact is a *contingent fact*, something that is a fact without being necessarily the case. And the temporal analogue of a subjective fact is a *temporary fact*, something that is a fact without being permanently the case. For a subjectivist like me, these analogies are both encouraging and useful. They are encouraging because if there's nothing incoherent in the notion of a contingent or a temporary fact (and there does not seem to be), chances are that the notion of a subjective fact is also one that we can make decent sense of. And they are useful because they allow us to formulate Subjectivism (the view that reality is only subjectively the way it is) in analogy to *Contingencism* (the view that reality is only contingently the way it is) and *Temporaneism* (the view that reality is only temporarily the way it is).[2]

So let us begin by taking a closer look at these two views, both of which are more popular and familiar than Subjectivism. While there are certainly various ways of formulating them, my preferred one involves talk of *propositions*. I use the term "propositions" to refer to whatever things are the objects of belief and other propositional attitudes and the semantic values of declarative sentences relative to contexts.[3] While I assume that there are entities of

2 I chose these neologisms because the terms "Temporaryism" and "Contingentism" have recently been used by Williamson (2013) to refer to certain controversial theses about ontology, while the term "temporalism" belongs to a long-standing debate in semantics (Richard 1980). "Temporaneism" is my label for the view that McTaggart (1908) called (somewhat unhelpfully) the "A-theory" of time. The analogy between subjectivity, modality and time is a central theme in the philosophy of Arthur Prior (see, in particular, Prior and Fine (1977)).

3 This is the way the term "propositions" has traditionally been used. There are, of course, dissenters – most notably, Lewis (1980) argued that the things that are objects of our attitudes are not also the semantic values of sentences relative to context. For a recent response, see Cappelen and Hawthorne (2009).

this kind, and that at least some of them can instantiate the monadic properties of truth simpliciter and falsity simpliciter, I make no assumption concerning their metaphysical nature (in particular, whether they are structured or unstructured, coarse- or fine-grained, sparse or abundant, etc.).

Contingencism is the thesis that reality is only contingently the way it is. But if reality is only contingently the way it is, some propositions that happen to describe reality correctly do not do so necessarily. This means that, while these propositions are true simpliciter, they do not hold true under all possible circumstances or, as it is usually put, *in all possible worlds*.[4] So the essence of Contingencism seems to be the claim that:

(*Contingencism*) Some propositions are true simpliciter without being true in all possible worlds.

From here to the notion of a contingent fact it's a small step. When a proposition is true simpliciter it can be said to express a fact (for what is truth *simpliciter* if not the property of reflecting the way reality is *in and of itself*?). But when it is not true under all possible circumstances, a proposition cannot be regarded as necessary. Hence the formulation of Contingencism I just offered carries with it a commitment to facts that are contingent rather than necessary, as was to be expected.

The negation of Contingencism is *Necessitarianism* – the view that reality is necessarily the way it is or, equivalently, that there are no contingent facts because every proposition true simpliciter is true in all possible worlds. A champion of Contingencism is Leibniz, according to whom there must be contingently true propositions, "otherwise

4 For present purposes, I will ignore the distinction, drawn by Adams (1981) between truth *in* a possible world and truth *at* a possible world.

everything would be necessary and nothing would be possible other than that which actually attains existence" (Leibniz 1989, 28). A champion of Necessitarianism is Spinoza, who thought that "nothing in nature is contingent", because "things could not have been produced by God in any other way or in any other order than is the case" (Spinoza 2002, 234-235).

On to Temporaneism. Temporaneism is the view that reality is only temporarily the way it is. But if reality is only temporarily the way it is, some propositions that happen to describe reality correctly do not *always* describe it correctly. This means that, while these propositions are true simpliciter they do not hold true *at all times*. So the essence of Temporaneism is the thesis that:

(*Temporaneism*) Some propositions are true simpliciter without being true at all times.

Once again, if every proposition that is true simpliciter expresses a fact, it's easy to see that this formulation of Temporaneism carries with it a commitment to the existence of temporary facts.

The negation of Temporaneism is *Sempiternalism*, the view that reality is eternally the way it is or, equivalently, that all facts are eternal because every proposition true simpliciter is also true at all times.[5] A champion of Temporaneism is Arthur Prior, who mocked Sempiternalism as the view that reality is "a timeless tapestry with everything stuck there for good and all" (1998, 104) . A champion of Sempiternalism is J. J. C. Smart, according to whom "the transitory aspect of time […] is an illusion that prevents us seeing the world as it really is" (1998, 94)

With this basic characterizations of Temporaneism and Contingencism in place, we

5 Sempiternalism should not be confused with Eternalism, the ontological view that past and future objects exist just as much as present ones. For a a discussion of the relationship between Sempiternalism (or the 'B-theory' of time, as it is often called) and Eternalism, see Zimmerman (2005).

can finally move on to consider Subjectivism. Just as contingencists think that how reality is is a contingent matter and temporaneists think that how reality is is a temporary matter, subjectivists think that how reality is is a subjective matter – they think that reality is only subjectively the way it is. But what does it mean to say that reality is only *subjectively* the way it is? At a first pass, it means that some propositions that happen to describe reality correctly do not *objectively* do so – subjectivity being to objectivity what contingency is to necessity and temporariness to sempiternity. But, given what we said about Temporaneism and Contingencism, the analysis can be pushed a little further. Necessary truth is standardly defined as truth *in all possible worlds* and sempiternal truth is standardly defined as truth *at all times*. Why not adopt a similar strategy here and think of objective truth as truth that holds across a series of 'points' analogous to (but, obviously, different in kind from) possible worlds and times? I suggest we call the points in question "points of view" and formulate Subjectivism as follows:

(*Subjectivism*) Some propositions are true simpliciter without being true from all points of view.

It is important to see that this claim wouldn't be of much interest if by "point of view" we meant what is ordinarily meant by this expression, i.e. someone's opinion about something or, quite literally, someone's perceptual perspective. For then Subjectivism would reduce to the platitude that some propositions, though true simpliciter, are contradicted by someone's opinions or by the way things look from someone's perceptual perspective. However, the analogy with possible worlds and times suggests a different way of using the expression "point of view". When we think of a possible world, we think of a way in which reality could manifest itself – a *possible manifestation* of reality, as one might put it. We do *not* think (or, at

least, we are reluctant to think) of possible worlds as ways in which reality can be *imagined* to be. For couldn't reality manifest itself in ways that no one is actually able to imagine? Analogously, when we think of a time – a past time, let's suppose – we do *not* think of it as a way in which reality can be *remembered* to be. For isn't there more to the past than anyone can possibly remember? A past time is just a way in which reality manifested itself in the past – a certain kind of *temporary manifestation* of reality. The situation is similar with points of view. Just as we need not reduce possible worlds to sets of imaginings and past times to sets of memories or recollections, we need not reduce points of view to sets of perceptions or opinions. Instead, we can think of a point of view as a way in which reality manifests itself to someone– a *subjective manifestation* of reality. In doing so, we allow that something may be the case from a certain subject's point of view without that subject (or anyone else, for that matter) taking notice of it, either perceptually or doxastically. It is this metaphysical notion of a point of view – which is, arguably, as basic as the notion of a possible world or a time – that subjectivists need in order to state their position: a subjectivist thinks that *reality in and of itself* varies across different subjects, not just that reality is *believed* or *perceived* to be different by different subjects. And this is just another way of saying that, for a subjectivist, at least some of the facts that constitute reality are subjective rather than objective.

The negation of Subjectivism is *Objectivism* – the thesis that reality is objectively the way it is or, equivalently, that there are no subjective facts because every proposition true simpliciter is also true from every point of view. An objectivist is someone who thinks that, while there may be subjectivity in how we apprehend or evaluate reality, there is no subjectivity in how reality is in and of itself: subjectivity may shape our perceptions and opinions, but it simply does not run as deep as to shape the facts. With a few exceptions, this is the dominant view in contemporary metaphysics.

It is important to see that, as I chose to formulate them, Contingencism,

Temporaneism and Subjectivism make essential use of the property of truth simpliciter. This differentiates them from three other theses about propositions, namely:

(*World-relativism*) Some propositions are true in some possible worlds without being true in all possible worlds.

(*Time-relativism*) Some propositions are true at some times without being true at all times.

(*Subject-relativism*) Some propositions are true from some points of view without being true from all points of view.

World-relativism, Time-relativism and Subject-relativism say nothing about what sort of things can be true simpliciter. So a good case can be made that they have no direct implication for whether the way reality is in and of itself is contingent, temporary or subjective. For example, a time-relativists who thinks that among the possible objects of belief and other propositional attitudes there is the time-relative (or 'temporal') proposition *that it is raining in Paris* need not believe that that proposition can be true (or false) simpliciter – she need not believe that there is any non-time-relative fact of the matter as to whether it is raining in Paris.[6] Similarly, a subject-relativist who thinks that among the possible objects of belief and other propositional attitudes there is the subject-relative (or 'subjective') proposition *that chocolate is tasty* need not believe that that proposition can be true (or false) simpliciter – she need not believe that there is any non-subject-relative fact of the matter as to whether chocolate is tasty.[7] What's distinctive of contingencists, temporaneists and subjectivists is that,

6 Mellor (1998) and Sider (2001) are both examples of this: they accept Time-relativism without accepting Temporaneism.

7 See, for instance, Kolbel (2003) and Lasersohn (2005). The difference between Subjectivism and Subject-

besides accepting the existence of world-, time- and subject-relative propositions, they also deem (some of) these propositions capable, all by themselves, of describing reality correctly or incorrectly and, therefore, of instantiating the monadic properties of truth simpliciter and falsity simpliciter.

Some will see a tension here. Isn't it contradictory to say that a proposition is true *simpliciter* and then go on to describe the very same proposition as true '*in*' some worlds and not others, '*at*' some times and not at others or '*from*' some points of view and not others? No, it isn't. Just as one can say that the proposition *that God does not exist* is true simpliciter and then go on to say that the very same proposition is not true *according to* the Bible, one can say that some propositions are true simpliciter and then go on to say that the same propositions are not true *in* some worlds, *at* some times and *from* some points of view. There are two kinds of properties at stake here: a monadic property (*being true*) and a bunch of relations (*being true in*, *being true at*, *being true from*). Contingencists, temporaneists and subjectivists need not choose between one kind of property and the other. At most, they owe us an explanation of how they are related. But this they can easily do.

Contingencists take what is true absolutely or simpliciter to be what is true in a particular world, the actual world:

A proposition is true simpliciter iff it is true in the actual world.

---

relativism is obscured by the fact that Subject-relativism licenses assertions of the form “It is a fact that chocolate is tasty” or “It is true that chocolate is tasty”. But the whole point is that, on any standard subject-relativist account, the very truth of these assertions is relativized to this or that point of view. The dialectic here is familiar: “The relativist [...] does indeed allow a syntactically monadic truth predicate that behaves in a disquotational way (roughly, 'S' is true relative to a parameter value iff "S' is true' is true relative to that value). But [he] does not think of 'true' as expressing a monadic property” (Cappelen and Hawthorne 2011, 460).

Of what is true in some possible world or another, contingencists say that it is *possibly* true:

A proposition is *possibly* true iff it is true in some possible world.

Temporaneist do something similar. They take what is true absolutely or simpliciter to be what is true at one particular time, the present time:

A proposition is true simpliciter iff it is true at the present time.

Of what is true at times that precede or follow the present time, temporaneists say that it is *was* or *will be* true:

A proposition *was* true iff it is true at a time earlier than the present time.

A proposition *will be* true iff it is true at a time later than the present time.

Subjectivists will adopt a similar strategy. They will identify a certain point of view – call it the "firstpersonal" point of view – such that all and only what is true from that point of view is also true simpliciter:

A proposition is true simpliciter iff it is true from the firstpersonal point of view.

They can then talk of what is true from any point of view other than the firstpersonal one as being only *otherpersonally* true:

A proposition is *otherpersonally* true iff it is true from a point of view other than the

firstpersonal one.

So, for example, a subjectivist who thinks that the proposition *that chocolate is tasty* is true simpliciter (i.e. that, contrary to what is generally assumed, there is an absolute fact of the matter as to the tastiness of chocolate) may concede that that proposition is otherpersonally false, meaning that it is not true from some points of view other than the firstpersonal one. Notice that, on the resulting picture, the proposition *that chocolate is tasty* does double duty: it reflects how reality is in and of itself, while *also* reflecting one point of view among others. Subjectivists achieve this combination (and thereby vindicate the notion of a subjective fact) because they identify the way reality is in and of itself with the way reality is according to one particular point of view – the point of view which they call "firstpersonal".

It is straightforward to see that, given the foregoing theses about the relationship between monadic and relative truth (and assuming that when a proposition is not true from a given standpoint, it is false from that standpoint), Contingencism, Temporaneism and Subjectivism can also be stated in the following ways:

(*Contingencism**) Some propositions that are true simpliciter are possibly false.

(*Temporaneism**) Some propositions that are true simpliciter were or will be false.

(*Subjectivism**) Some propositions that are true simpliciter are otherpersonally false.

Some will prefer these formulations to the ones I gave earlier on the ground that they are formulated in terms of the notion of monadic truth, which they take to be conceptually more basic. I myself do not want to take a stand on this issue (for example, I think one can define the notion of the actual world in terms of monadic truth, but I have no principled objection to proceeding the other way around). Others will prefer these formulations to the ones I gave

earlier on the ground that they do not overtly quantify over such things as possible worlds, times and points of view. But even if they do not overtly quantify over these things, it is unclear that they avoid ontological commitment to them – at any rate, I am not going to assume that they do. In what follows, I will switch back and forth freely between talk of something being otherpersonally true and talk of something being true from some point of view other than the firstpersonal one – nothing of what I will say hinges crucially on whether or not points of view can be 'paraphrased away' in terms of the notion of what is otherpersonally the case. I will also switch back and forth freely between talk of the proposition that p being true simpliciter but otherpersonally false and talk of the fact that p being a subjective fact – for my purposes, there is no need to choose between an ontology of propositions and an ontology of facts.

2. The World As I Found It

Subjectivism is the view that some propositions are true simpliciter without being true from all points of view. In the last section, I defended the intelligibility of this view. But why do I believe this view to be true? In a nutshell, because I happen have to certain intuitions about myself and my place in reality and I think that these intuitions cannot be vindicated if Subjectivism is false. This answer will not satisfy everyone. In particular, it will not satisfy those who deny that intuitions can, all by themselves, provide one with good *reasons* for believing this or that philosophical view. I disagree with this claim, but a defence of the epistemic role of intuitions in philosophical theorizing would take me too far afield. Holders of the view that intuitions are epistemically idle can take what I am going to say in this section as an account (or, if they want, a confession) of how I came to believe Subjectivism in

the first place. In the second part of this thesis, when I will look at some concrete applications of the Subjectivist View of the Mental, I will try to show that the attraction of Subjectivism goes well beyond its intuitive appeal. I present intuitions first simply because they are what first made me see Subjectivism in a favourable light.

What are these intuitions that I have? Really, they boil down to a very simple thought: that, of all individuals in the world, the individual I am, Giovanni, is somehow *special*. Put it this way: if I were to write a book entitled "The World As I Found It" or "The World As It Really Is", Giovanni would have a role in that book that no other individual has. He would be (I blush to say) the *main character* of that book, the only and authentic *center* of the world. That, of all individuals there are, Giovanni is the one having this role strikes me as an undeniable and all-too-important fact. To me, writing the book of the world without mentioning the fact that Giovanni is special would be writing an incomplete book.

The intuition will sound very vague, but it can be made more precise. In what sense is Giovanni special? What does this specialness consist in? To be sure, I don't want to deny that Giovanni resembles other subjects in many respects. For example, just like any other subject, Giovanni has mental states of various sorts: fears, desires, beliefs, hopes, thoughts, experiences and feelings. But then, again, take Giovanni's fears. It seems to me to be a fact as clear as daylight that what Giovanni fears is more quintessentially *fearsome* than what other people fear. Or take Giovanni's desires. It seems to me to be a fact as clear as daylight that nothing is as quintessentially *desirable* as what Giovanni desires. Of all fears and all desires, Giovanni's fears and desires are (if I may use a typographical trick to convey this intuitive point) FEARS and DESIRES: they are fears and desires *par excellence* because they make their objects truly and quintessentially fearsome and truly and quintessentially desirable.

What goes for fears and desires goes also for other intentional states. What Giovanni believes is, ipso facto, more *credible* than what other subjects believe. If Giovanni is

interested in a certain thing, this makes that thing far more *interesting* than if other people are interested in it. When Giovanni hopes that this or that will happen, it very *important* whether this or that will happen – more important than whether other people's hopes will be fulfilled or not. In short, just as Giovanni's fears and desires are FEARS and DESIRES, Giovanni's beliefs, interests and hopes are BELIEFS, INTERESTS and HOPES.

And what goes for intentional states goes also for non-intentional states (or states that have been alleged to be non-intentional). I would say that there's nothing as *painful* as Giovanni's pains, nor anything as *pleasant* as Giovanni's pleasures. More in general, I would say that Giovanni's experiences are quintessentially *experiential* and that Giovanni's feelings *make themselves felt* in a way in which the feelings of no other subjects do. Giovanni's pains, pleasures, experiences and feelings are PAINS, PLEASURES, EXPERIENCES and FEELINGS.

To generalize, Giovanni's mental states, and only Giovanni's mental states have some unmistakeable 'glow' to them that makes them MENTAL states. And the difference between MENTAL states and the mental states of others couldn't be starker. To borrow the words of William James:

> […] the former have a warmth and intimacy about them of which the latter are completely devoid, being merely conceived, in a cold and foreign fashion, and not appearing as blood-relatives, bringing their greetings to us from out of the past. (James 1950, 332).

Sceptics try to convince me all the time that all this is nonsense. “There's nothing special about Giovanni” – they say – “except that *you* are Giovanni. And there's nothing special about Giovanni's mental states except that they are *your* mental states. If that is all you are aiming at, there is not much metaphysical substance to the intuitions you are appealing to”.

Not so quickly. Suppose the (alleged) specialness of Giovanni's mental states reduced

to the fact that they are *my* mental states. Then presumably the same asymmetry I (seem to) observe between my mental states and everyone else's mental states I would also (seem to) observe between *my* shoes and everyone else's shoes or between *my* nose and everyone else's nose. And so I would be tempted to distinguish between shoes and SHOES and between noses and NOSES just as much as I am tempted to distinguish between mental states and MENTAL states. But of course I am not tempted to do that: my intuitions tell me that my mental states are special, while my shoes and nose are not. So, whatever it consists it, the specialness of my mental states cannot be reduced to the fact that they are my mental states.

"Fair enough. Giovanni's mental states are special in a way that Giovanni's shoes or Giovanni's nose are not. But they are special *for Giovanni*, in the same way as Mary's mental states are special *for Mary* and Fred's mental states are special *for Fred*. We all bear a special relationship to our own mental states – a relationship that we do not bear to our nose or our shoes. So what? Haven't we known this all along?"

This reply misunderstands the import of my intuitions. It may be true that all mental states have the property of being special for their owner and nobody else. In this respect, they may well be all on a par. But the point of my intuitions is precisely that, no matter what properties are shared across all mental states, there must *also* be some property that distinguishes my mental states from the mental states of others. As Wittgenstein's interlocutor says in this passage from the *Philosophical Investigations*:

> 398. "When I imagine something, or even actually *see* objects, I have *got* something which my neighbour has not." – I understand you. You want to look about you and say: "At any rate only I have got THIS." (Wittgenstein 1986, 120)

I want to look about and say that that there is a property that only *my* mental states have. If all mental states have the property of being special *for* their owner, then perhaps my mental states

have the property of being special, *full stop*. That's what I am referring to when I say that, unlike all other mental states, my mental states are MENTAL states.

"Fair enough, it *seems* to you as if your mental states were different from the mental states of others. But that's just because you know them *directly*, via introspection, and you know the mental states of others *indirectly*, based on the observation of their behaviour. Your mental life is *not* different from all other mental lives, it's just that you constantly look at it through the deforming lens of introspection".

I think I could easily accept this error-theory if I could easily accept the idea of introspection. But, deep down, my intuitions recalcitrate. It is not so much that I agree with Ryle that "'introspection' is a term of art and one for which little use is found in the self-descriptions of untheoretical people" (Ryle 1949, 152). And it is not so much that I trust the authority of those psychologists who say that self-knowledge and knowledge of other minds are *not* associated with radically different ways of knowing.[8] It is just that if I pay close attention to what I do when I try to find out how things are with my mental life, I find no unitary method or process – no specific act of 'looking within' that deserves the name of "introspection".[9] Phenomenologically speaking, the deforming lens of introspection is nowhere to be found. And if there is no deforming lens, there can't be any deformation: that my mental states are MENTAL states is not a *illusion*. It is simply a *fact*, as hard a fact as any other.

"You may reject the notion of introspection. But you certainly won't deny that you know your own mental states *better* than you know other people's mental states. That's the source of your mistake: you take your mental states to be *metaphysically* privileged just because you happen to have some kind of privileged *access* to them".

---

8 See Carruthers (2011).

9 See Schwitzgebel (2012) for a defence of this point with which I am very much in agreement. I will say much more about introspection in Chapter 3.

I am not sure I have privileged access to my mental states. But whether I do or not is quite beside the point here. I could have the most extensive knowledge of other people's mental states – I am sure that that wouldn't make them MENTAL states. And I could have the poorest and most superficial knowledge of my own mental states – I am sure they would be MENTAL states nonetheless. My intuitions have nothing to do with how well I *know* my mental life. They have to do with how my mental life *is*, in and of itself.

"So you really believe this claim – that Giovanni is special and that his mental states are MENTAL states – to express an objective fact of the matter about how reality is in and of itself?"

Two separate questions are packed together here. Do I believe the claim that Giovanni is special to reflect the way reality is in and of itself? Yes, I do, otherwise I wouldn't call it a "fact". But do I believe it to be an *objective* fact, the kind of fact that obtains *from all points of view*? Of course not. My intuitions tell me that Giovanni is special, but they also tell me that, *from some other point of view*, Giovanni is not special and other individuals are special instead. Take you, for example. If, after reading what I've written so far, you decided to trust the same kind of intuitions I decided to trust, you would come to believe that *you* are special and that *your* mental states are MENTAL states. It's impossible for me not to acknowledge that *from some other point of view* – i.e. according to the way reality manifests itself to some individual other than myself – things are, indeed, as those beliefs represent them to be.[10] And acknowledging this is acknowledging that there's a point of view from which you are special, and Giovanni is not. That Giovanni is special *is* a fact. But it is a subjective fact, rather than an objective one.

To the extent that I trust these intuitions, the picture of reality I am drawn to looks

10 In the next chapter, I will explain how this observation can be used to vindicate the correctness of your belief that you are special and that your mental states are MENTAL states.

more or less like this:

> Reality contains many things: mountains and rivers, plants and animals, stars and planets. Facts concerning these things are all objective: they obtain from every point of view. Among the many things reality contains, there are individuals who enjoy mental states: beliefs, hopes, desires, feelings and the like. Facts concerning which individuals there are and what kind of mental states they enjoy are also objective facts. But then, alongside these objective facts, there are some subjective facts, too. Chief among them, the fact that Giovanni is special, i.e. that his beliefs, hopes, desires, and feelings are BELIEFS, HOPES, DESIRES and FEELINGS. This fact is subjective because there are points of view from which it does not obtain – points of view from which it is someone else's mental states that are MENTAL states, and not Giovanni's.

The view embodied in this picture is what I call the "Subjectivist View of the Mental" (SVM, hereafter). If I were to write a book entitled "The World As I Found It" or "The World As It Really Is", that book would be a detailed and fully worked-out version of SVM. And it's because I've been intrigued by SVM that I came to embrace Subjectivism, the general doctrine that reality is only subjectively the way it is.

3. The Subjectivist View of the Mental

In formulating SVM, I said that the fact that Giovanni's mental states are MENTAL states is a subjective fact. Another way of expressing the same point would be to say that the

proposition *that Giovanni's mental states are* MENTAL *states*, though true simpliciter, is otherpersonally false.

These formulations are – I take it – reasonably clear, but they are also somewhat cumbersome. Reference to entities like facts and propositions was useful to arrive at a clear statement of Subjectivism. But, at this point, it would be convenient if we could develop a formal language in which the central claims of SVM – claims about what is true, false, otherpersonally true or otherpersonally false – can be expressed in a more direct fashion. For instance, instead of using a certain sentence *s* to construct a noun-phrase referring to a proposition ("the proposition that *s*") and say of that proposition that it is otherpersonally true, it would be nice if we could simply assert another sentence *s**, such that for the proposition expressed by *s** to be true is for the proposition expressed by *s* to be otherpersonally true. In the formal language in question, one would expect there to be a straighforward syntactic rule to generate *s** from *s*, as well as a notation making the relation between *s** and *s* fully perspicuous. The question is: how do we go about developing a language of this kind?

One obvious possibility is to look at how the same problem has been dealt with in a different, but closely related domain. Suppose you are a temporaneist and you are interested in talking about what was the case (but may or may not be presently the case) without mentioning facts or propositions. One way you could do this is by enriching your language with a sentential operator that functions in roughly the following way: whenever prefixed to a sentence *s* it generates another sentence *s** that is true if and only if the original sentence *s* was true. The advantage of having this operator at your disposal is that, instead of having to assert or deny the past truth of the proposition that *s*, you could simply assert or deny *s**.

It's not clear that English contains operators of this kind. There is the locution "It was the case that...", but the result of prefixing it to, e.g., "It is raining in Paris" is a sentence of dubious grammaticality and meaning. We can, of course, obtain the new sentence by

modifying the tense of the verb in the old one. But, pending an account of how tense works in natural language, it would be incautious to assume that the relation between the proposition expressed by “It is raining in Paris” and the proposition expressed by “It was raining in Paris” is as straightforward as temporaneists would it like it to be.[11]

English may not contain the operators temporaneists want, but there is a formal language that does. It's a language developed by Arthur Prior in the late 1950s and early 1960s and generally referred to as “Temporal Logic”. The simplest version of propositional Temporal Logic is a logic that, alongside the familiar truth-functional connectives ('~' for negation, '&' for conjunction, 'v' for disjunction, '→' for material implication), contains two non-truth-functional sentential operators, 'WAS' and 'WILL'. The rule governing their meaning is simple: if you already know the truth-value of a certain sentence *s* relative to (or 'at') each time, you can determine the truth value of the sentences obtained by prefixing 'WAS' or 'WILL' to *s* by looking at the past and future truth-values of *s*. A bit more formally:[12]

(1) *WILL s* is true at a time t iff there is some time t’ later than t and *s* is true at t’

(2) *WAS s* is true at a time t iff there is some time t’ earlier than t and *s* is true at t’

The syntactic structure of *WAS s* and *WILL s* reminds us of the semantic relation these sentences bear to the unadorned *s*. This pleasing 'correspondence' between syntax and semantics – a correspondence we seldom observe in the sentences of English and other natural languages – is one of the main reasons why temporaneists find Temporal Logic so congenial to their purposes.

11 For a discussion of the semantics of tense in natural language, see King (2003).

12 Hereafter, in order to avoid proliferation of quotation marks, I will be sloppy about the distinction between use and mention.

If we are to model the official idiom of Subjectivism on Temporal Logic, we need a sentential operator doing for us what 'WAS' and 'WILL' do for a temporaneist. Let the new operator be 'OTHER' (to be read “otherpersonally”). In Temporal Logic, we define the truth-value of *WAS s* and *WILL s* on the basis of the truth-value that *s* has relative to past or future times. Here we are going to define the the truth-value of *OTHER s* in terms of the truth-value that *s* has relative to other points of view. Our starting point will be a model including a set of objects called “points of view” and an interpretation function assigning truth or falsity to each atomic sentence relative to each point of view. When it comes to sentences containing only the usual truth-functional connectives, we are going to say that their truth-value relative to a certain point of view is determined – in the usual way – on the basis of the truth-value, relative to the same point of view, of the atomic sentences they contain. But when it comes to sentences containing 'OTHER', we are going to say that:

(3) *OTHER s* is true from a point of view v iff there is some point of view v' distinct from v and *s* is true at v’

Using OTHER, we can define two operators that will come in useful later on, 'SOME' (to be read as “somepersonally”) and 'ALL' (to be read as “omnipersonally”):

$$\text{SOME } p =_{def} p \vee \text{OTHER } p$$

$$\text{ALL } p =_{def} \sim \text{SOME} \sim p$$

A propositional logic enriched with these devices – or propositional 'Firtpersonal Logic', as I shall call it – gives the subjectivist what she wanted: a method from constructing, from any sentence *s*, a sentence *s** such that for *s** to be true is for *s* to be otherpersonally true. That

sentence is, of course, *OTHER s* and has the appealing feature of wearing its meaning on its sleeves.

Needless to say, propositional Firstpersonal Logic shares the limitation of any other propositional logic. To express many of the claims they want to express and draw all the distinctions they want to draw, subjectivists will need to construct a *quantified* Firstpersonal Logic. The language of quantified Firtpersonal Logic is just the language of predicate logic (including predicates, variables and quantifiers binding them), with 'OTHER' added. Unsurprisingly, quantified Firstpersonal Logic is to propositional Firstpersonal Logic what quantified Temporal Logic is to propositional Temporal Logic. If quantified Temporal Logic contains sentences of the form:

WAS ∃x Fx

∃x WAS Fx

Quantified Firstpersonal Logic will contain sentences like:

OTHER ∃x Fx

∃x OTHER Fx

which can be translated in English as:

Otherpersonally, there is an x which is F

There is an x which is otherpersonally F.

The structural similarities between the two languages should, at this point, be apparent. A

semantics for quantified Firstpersonal Logic can be provided in terms of points of view and point-relative truth in exactly the same way in which a semantics for quantified Temporal Logic can be given in terms of times and time-relative truth: since the latter is both straightforward and well-known, I will leave it to the reader to draw the obvious parallels.

What interests me here is to use quantified Firstpersonal Logic to offer a precise formulation of the main tenets of SVM. For convenience of presentation, I will have my variables and quantifiers range solely over subjects (where these are thought of as individuals who have mental states of some sort). I will assume that subjects exist objectively, i.e. that from every point of view every subject exists and is a subject from any other point of view. And I will adopt the convention of calling a particular subject *special* ('Sx') if and only if some of that subject's mental states are MENTAL states. My aim will be to state what SVM says from the firstpersonal point of view, which is also what SVM says simpliciter. At this stage, no assumption should be made about what SVM says from points of view other than the firstpersonal one. In particular, the reader *not* assume that the content of SVM is invariant across different points of view. In the next chapter, I explain why, given the truth of Subjectivism, it is best to understand SVM as a theory that, from any point of view other than the firstpersonal one, says about someone else what, from the firstpersonal point of view, it says about Giovanni.

The central claim of SVM is that there is one and only one subject who is special and that subject is Giovanni. This can be said without using any particular operator:

$$(S1)\ \exists x\ (x = \text{Giovanni} \ \&\ Sx \ \&\ \forall y\ (Sy \rightarrow y{=}x))$$

The second central claim of SVM is that for every subject there is some point of view from which that subject and only that subject is special – in a slogan, “each of us gets a subjective

chance to be the special one".[13] This can be put by saying that, from every point of view, every subject x is such that, somepersonally, x is special and no one else:

(S2) ALL ∀x SOME (Sx & ∀y (Sy → y=x))

The third central claim of SVM is that for every subject there is just *one* point of view from which that subject is special – with another slogan, "each of us gets only one subjective chance to be the special one".[14] To express this, we need to say that, from every point of view, if someone is special, that person is not also otherpersonally special:

(S3) ALL ∀x (Sx → ~ OTHER Sx)

From these three principles, one can deduce several claims that were already part of the informal presentation of SVM I offered at the end of the last section. For example, these three principles entail that I am not otherpersonally special (i.e. that ~ OTHER S (Giovanni)): if I were otherpersonally special, there would be more than one point of view from which I am special, contrary to what S3 dictates. They entail that all subjects except Giovanni are otherpersonally special (i.e. that ∀x (x ≠ Giovanni → OTHER Sx)): if they were not otherpersonally special, there would be no point of view from which they are special, contrary to what S2 requires. They also allow us to define, for every subject x, what it means for such-and-such to be the case "from x's point of view": since (by S2 and S3) there is just one point of view from which x is special, we can take "from x's point of view, p" to mean that,

13 One may think of this claim as the subjective analogue of the claim that every instant is sometimes the only present instant.

14 One may think of this claim as the subjective analogue of the claim that every instant is only ever present *once.*

omnipersonally, p is the case if x is special. In symbols:

$$\text{From x's point of view, } p =_{def} \text{ALL } (Sx \rightarrow p)$$

S1, S2 and S3 tell us a lot about SVM, but they still do not tell us everything. In particular, they do not tell us much about the relationship between MENTAL states (BELIEFS, HOPES, DESIRES, EXPERIENCES, etc.) and mental states (beliefs, hopes, desires, experiences, etc.). We know that some mental states (e.g Giovanni's mental states) are also MENTAL states, while other mental states (e.g. the mental states of subjects other than Giovanni) are otherpersonally MENTAL states. What we do not yet know – and S1, S2 and S3 do not allow us to know – is whether these correlations are just a coincidence or, rather, the result of some kind of deeper connection between the mental and the MENTAL.

The last claim I want to incorporate into SVM addresses precisely this question. It says that it lies in the very nature (or essence) of a mental state to be *either* a MENTAL state *or* an otherpersonal MENTAL state. The idea is that mental states are – essentially and not just accidentally – a mixed bag: in much the same way in which it lies in the nature of jade to be *either* jadeite *or* nephrite, it lies in the nature of a belief to be *either* a BELIEF *or* an otherpersonal BELIEF and in the nature of a pain to be *either* a PAIN *or* an otherpersonal PAIN. That's why nobody can believe something without either BELIEVING it or otherpersonally BELIEVING it and nobody can be in pain without either being in PAIN or being otherpersonally in PAIN. The mental is only the surface – below the surface, there are MENTAL similarities and MENTAL differences and it is in virtue of these similarities and differences that any ordinary mental fact holds.

If we use '$\leftrightarrow_R$' to indicate the giving of a *real definition* – a definition specifying the metaphysical nature of what appears on its left hand side in terms of what appears on its right

hand side – we can put this point, more formally, as follows:

believe (x) $\leftrightarrow_R$ BELIEVE (x) v OTHER BELIEVE (x)

("To believe is to *either* BELIEVE *or* otherpersonally BELIEVE")

pain (x) $\leftrightarrow_R$ PAIN (x) v OTHER PAIN (x)

("To be in pain is to *either* be in PAIN *or* be otherpersonally in PAIN")

These claims can be seen as instances of a general schema relating mental states ('M') and MENTAL states ('$\mathbb{M}$'), and it is all the instances of this schema that will have to be incorporated into SVM:

(S4) Mx $\leftrightarrow_R$ $\mathbb{M}$ x v OTHER $\mathbb{M}$ x

The importance of adding such a general principle to the view will become clearer and clearer as we proceed. For the moment being, I will note just two things.

First of all, S4 allows us to be more specific about what it takes for a subject to be special. On my terminology, a subject is "special" if and only if some of that subject's mental states are MENTAL states. But S3 and S4 together imply that if *some* of a subject's mental states are MENTAL states then *all* of that subject's mental states are MENTAL state. For S4 tells us that any of that subject's mental state are either MENTAL states or otherpersonal MENTAL states. And S3 tells us that, if some of that subject's mental states are MENTAL states, none of that subject's mental states are otherpersonal MENTAL states (if they were, there would be more than one point of view from which that subject is special, contrary to what S3 requires). So, given S3

and S4, any special subject is a subject *all* of whose mental states are MENTAL states.

Secondly, S4 can be used to elucidate and clarify the notion of a MENTAL state. I introduced this notion in § 2 and the way I did it was by using certain metaphors (I said that the MENTAL states are those that 'glow' vis-à-vis all others), by inviting my reader to consider some concrete examples of MENTAL states (from one's point of view, the best examples of MENTAL states are one's own mental states), and by quoting philosophers who seem to be after the same distinction that I am after when I talk about mental states and MENTAL states. All this was supposed to help the reader get a grip on a notion that I was basically treating as primitive. But maybe this notion does not *have* to be treated as primitive. For, if we have S4, we can explicate the notion of a MENTAL state in terms of two other notions – the ordinary notion of a mental state and the notion of something being otherpersonally the case. For example, we can characterize PAIN as that state X such that to be in pain (in the ordinary sense) is to either be in X or be otherpersonally in X and we can characterize BELIEF as that state Y such that to believe something (in the ordinary sense) is to either be in Y or be otherpersonally in Y. Notice that these characterizations do *not* reflect what a defender of SVM like me takes to be the metaphysical order of priority (they result from reading S4 right-to-left, instead of left-to-right). But this is quite irrelevant. It's true that, for me, MENTAL states are metaphysically more fundamental than mental states. Nevertheless, I am perfectly happy to allow those of my readers who did not understand the metaphors, the examples and the quotes I used in § 2 to grasp the meaning of "PAIN" or "BELIEF" in terms of concepts that they are more familiar with, namely the ordinary concept of "pain" and the ordinary concept of "belief". This might be another one of those cases in which – as philosophers like to say – the metaphysical order (i.e. the order of being) and the conceptual order (i.e. the order of understanding) need not coincide.

4. Subjectivism: relatives and ancestors

So far, I have explained what I mean by "subjective fact", what intuitions lead me to believe in the existence of subjective facts and what kind of subjectivist view I endorse. But where does the notion of a subjective fact come from? Has Subjectivism ever been defended in the past? And if so, how do other versions of Subjectivism differ from SVM?

As far as I can tell, the first philosophers whose position has clear subjectivist undertones is G.W. Leibniz. Leibniz believed in the existence of a multitude (in fact, an infinity) of entities called "monads" that he described as "metaphysical points" (Leibniz 1989, 142). At the center of his metaphysics, there is the idea that the actual world (i.e. "the universe") literally varies across these different points:

> Just as the same city viewed from different directions appears entirely different and, as it were, multiplied perspectively, in just the same way it happens that, because of the infinite multitude of simple substances, there are, as it were, just as many different universes, which are, nevertheless, only perspectives on a single one, corresponding to the different points of view of each monad. (Leibniz 1989, 220)

Leibniz started from the assumption that every substance or monad is a "living mirror that represents the universe according to its own point of view" (Leibniz 1989, 211). In principle, this assumption would be compatible with the view that the universe, though represented differently by different monads, is objectively the way it is. But Leibniz did not content himself with this view and claimed, instead, that each substance *is* a universe or a concentrated world. It does not seem too much of a stretch to regard this position as a form of Subjectivism: according to Leibniz, the universe cannot be reduced to what is objectively the case, and since what is subjectively the case varies from the point of view of one substance to

that of another, what must be said is that "the universe is in some way multiplied as many times as there are substances" (Leibniz 1989, 42). Of course this 'multiplication' follows certain rules: the overall system is so harmonious that what is the case from one point of view can be read off what is the case from any other point of view (that's why, in the passage above, Leibniz is keen to emphasize that the different universes he talks about "are, nevertheless, only perspectives on a single one").[15] But the basic point remains that monads shouldn't be thought of as inhabitants of the same objective world. For Leibniz, "each monad is, as it were, a world apart" (Leibniz 1989, 206).

Leibniz never drew an explicit parallel between points of view, possible worlds and times (he did say that each substance is "a world apart", but he never likened the thesis that the actual world is 'multiplied perspectively' to the thesis that, alongside the actual world, there are other possible worlds). It was Arthur Prior who brought the analogy between worlds, times and subjects to everyone's attention. His 1968 paper "Egocentric Logic" discusses a formal language in which ordinary individuals play the same role that times play in Temporal Logic. In Temporal Logic – Prior noticed – atomic sentences are naturally thought of as sentences in the present tense – sentences that (at least in their unembedded occurrences) "are understood as directly or indirectly characterising the *un*mentioned time of utterance" (Prior 1968b, 193).[16] Obviously, this does not mean that Temporal Logic does not have the resources to talk about past or future times. The point is just that, in order to do so, one has to prefix the atomic sentences of Temporal Logic with operators like 'WAS' or 'WILL'. From a purely formal perspective, Prior's Egocentric Logic works in a similar way. Here atomic sentences (at

15 "Each substance expresses the whole series of the universe according to the point of view or relation proper to it, from which it happens that they agree perfectly" (Leibniz 1989, 76).

16 Prior talks also of tensed sentences "implicitly referring" to the time of utterance (ibid.). I take it that (whichever way we interpret the expressions "implicitly referring" and "characterizing") Prior's tensed sentences should *not* be thought of as containing a term referring to the time of utterance or a hidden argument-place for times.

least in their unembedded occurrences) are naturally understood “as directly or indirectly characterising the speaker” (Prior, ibid.) and sentential operators are prefixed to atomic sentences whenever the goal is to talk about subjects other than oneself. Curiously, when it came to inventing the operators in question, Prior took inspiration from Leibniz. He assumed, with Leibniz, that all individuals are arranged in a scale of comparative perfection and he relied on this assumption in explaining the meaning of the egocentric analogues of 'WAS' and 'WILL' – namely, 'Someone-more-perfect' (to be read as “from the point of view of a more perfect individual”) and 'Someone-less-perfect' (to be read as “from the point of view of a less perfect individual”). Thus, where Temporal Logic has *WAS p* and *WILL p*, Prior's Egocentric Logic has *Someone-more-perfect p* and *Someone-less-perfect p*.

As the reader will have noticed, there are important analogies between Egocentric Logic and the kind Firstpersonal Logic I described in the last section. But there are also important differences, and they do not have to do only with the vaguely Leibnizian foundations of Egocentric Logic. One crucial thing to note is that, in quantified Firstpersonal Logic, one can *both* refer to various individuals (using proper names and variables) *and* describe what is the case from the point of view of those individuals (chiefly, using 'OTHER'). This is not so on any version of Egocentric Logic. According to Prior, the whole point of Egocentric Logic was to do away with reference to and quantification over ordinary individuals. In Egocentric Logic, one cannot say such things as “An individual more perfect than me is drinking”. Instead one has to say “Drinking is the case from the point of view of a more perfect subject” (or, more appropriately, “Someone-more-perfect drinking”). The predicates of common speech – predicates true of some individuals and false of others – are replaced by egocentric sentences – sentences true relative to some points of view and false relative to others.

Interestingly, Prior was convinced that an analogous point would apply to Temporal

Logic. Any full-blown version of Temporal Logic – he thought – would *not* allow us to refer to or quantify over times, but only to characterize them indirectly, by means of various kinds of tense operators. Since he refused to think of times as *sui generis* entities and preferred to regard them as "logical constructions out of tensed facts" (Prior 1968b, 200), Prior took this to be an advantage rather than a limitation of Temporal Logic. But his attitude towards Egocentric Logic was completely different:

> Egocentric logic is a different matter; I find it hard to believe that individuals really are just [...] "points of view," or that the real world of individuals is just a logical construction out of such points of view. Nevertheless the fact that we can have a consistent and comprehensive egocentric logic as well as a logic of tenses does suggest that some sort of idealism or relativism is a more defensible philosophical position than it once looked. (1968b, 200).

For Prior, there was at least logical room for a coherent philosophical position that treated subjectivity in the way Temporal Logic (on his reading of it) invites us to treat time. But he equated this position to "some sort of idealism or relativism" and found it "hard to believe". As a consequence, he never took Egocentric Logic seriously, if not as a potential key to Leibniz's philosophy.

To my knowledge, the first metaphysician who took seriously the analogy suggested by Prior between time and subjectivity was Kit Fine. In his 2005 paper "Tense and Reality", Fine defines two metaphysical views, called respectively *Tense Realism* and *First-personal Realism*. Tense Realism is the thesis that "reality is constituted (at least, in part) by tensed facts" (2005, 271), the kind of facts that are expressed by tensed statements and that reflect the standpoint of the present time. First-personal Realism is the thesis that "reality is not exhausted by the 'objective' or impersonal facts but also includes facts that reflect a first-

person point of view" (2005, 311). In his paper, Fine is not concerned to defend any particular version of these views, but just to see how they might be best developed. His conclusion is that, if one accepts either Tense Realism or First-personal Realism, one should also accept a certain radical view called 'Fragmentalism' according to which "reality may be irredeemably incoherent" (Fine 2005, 281), i.e. constituted by facts with incompatible content.

On the face of it, there might not seem to be much of a difference between Tense Realism and Temporaneism or between First-personal Realism and Subjectivism. But, in fact, the differences exist and are significant. The first thing to be noticed is that two distinct senses of "reality" are involved in the formulation of these views. I have been using the term "reality" in what I take to be its ordinary sense: for me, for a proposition to reflect how things are in reality is simply for it to be true simpliciter. Fine, by contrast, employs a distinctively metaphysical concept of "reality". He thinks one can elucidate this concept using other distinctively metaphysical notions (such as the notion of something being "factual" and the notion of something being "grounded" in something else (Fine 2001)), but he doubts that one can define it (Fine 2005, 267) and he is adamant that a proposition's being true simpliciter is *not* enough for it to count as true "in reality" in his sense of the term (ibid.). To avoid confusion, then, let us call "metaphysical reality" or, more concisely, "meta-reality" what Fine calls "reality". Reality and meta-reality are two different things and, while Tense Realism and First-personal Realism are concerned with meta-reality, Temporaneism and Subjectivism are concerned with plain reality.

Perhaps this difference in subject matter would not be too important if what Temporaneism and Subjectivism say about reality matched exactly what, respectively, Tense Realism and First-personal Realism say about meta-reality. But this is not so. Temporaneism says that reality changes across time: what is now true simpliciter was or will not be true simpliciter. Subjectivism says that reality varies from one point of view to another: what is

true simpliciter may not be true from all points of view. Tense Realism and First-personal Realism do *not* make similar claims about meta-reality. In fact, the 'fragmentalist' view that, according to Fine, tense realists and first-personal realists should adopt is one on which meta-reality does *not* change over time and does *not* vary across subjects. Think of it this way: on Fine's preferred version of Tense Realism, meta-reality is constituted by a certain collection of facts and, although some of these facts can be characterized as temporary or 'tensed', the collection itself remains exactly the same as time passes – every fact that constitutes meta-reality now, whether 'tensed' or not, has always done and will always do so. Similarly, on Fine's preferred version of First-personal Realism, meta-reality is constituted by a certain collection of facts and, while it's true that some of these facts reflect a 'first-personal' point of view, it's also true that the collection itself is the same from every point of view – every fact that constitutes meta-reality, whether 'first-personal' or not, does so in a perfectly objective fashion.

All this being so, I am reluctant to draw any direct connection between First-personal Realism and the view I defend in this thesis. I do not think that First-personal Realism entails Subjectivism (the latter requires reality to vary across subjects, the former doesn't). Nor do I think that Subjectivism entails First-personal Realism (the former is compatible with there being no distinction between reality and meta-reality, the latter isn't). I do think that subjectivists can make fruitful use of Fine's notion of meta-reality. But, even with respect to meta-reality, my conclusions will not be the same as Fine's. According to Fine, first-personal realists are better off saying that meta-reality is constituted by facts with incompatible contents. According to me, subjectivists are better off saying that meta-reality – conceived of as the domain of what "grounds" or metaphysically explains what is the case in reality – is not constituted exclusively by facts. All these delicate points will be clarified and discussed at length in the next chapter.

Arthur Prior was not a subjectivist and neither is Kit Fine. Is there anyone, among contemporary philosophers, who explicitly embraced Subjectivism? Yes, there is. Caspar Hare, in his interesting book *On Myself and Other Less Important Subjects* (2009), advocates a form of Subjectivism that he calls "Egocentric Presentism" (EP, hereafter). If I were an Egocentric Presentist instead of a defender of SVM, I would endorse a picture along these lines:

> Reality contains many things: mountains and rivers, plants and animals, stars and planets, subject and mental states. Truths concerning these things are almost all objective. What is subjective is the fact that all and only the perceptual objects of Author's mental states (e.g. the table he is looking at, the itch he feels in his neck, etc.) instantiate *presence*. This fact is subjective because there are points of view from which it does not obtain – points of view from which it is the perceptual objects of someone else's mental states that instantiate presence.

How does EP differ from SVM and why do I prefer the latter to the former? Subtleties aside, the central difference regards the nature of subjective facts: Hare and I agree that reality (in the ordinary sense of the term) contains subjective facts, but we disagree on what kind of facts they are. According to Hare, subjective facts concern the distribution of a certain property he calls 'presence' (hence the label of 'Egocentric *Presentism*'). According to me, subjective facts concern the distribution of MENTAL properties (hence the label of 'Subjectivist View *of the Mental*').

Now, one reason why I prefer SVM to EP has to do with what Hare says about presence. He says that presence is instantiated by perceptual objects (Hare 2009, 21-22) and

that perceptual objects include both mental particulars (pains, itches, etc.) and garden-variety macroscopic objects (telephones, cars, paintings etc.). Though not completely counterintuitive, this setup raises questions that I would have a hard time answering if I were an Egocentric Presentist. For example, if the painting I am looking at instantiates presence, do all of the painting's parts instantiate presence (including, e.g., the back surface of the painting or the single atoms that the painting is composed of)?[17] And if I am observing a star explosion through a telescope, does that star explosion instantiate presence *now* despite having occurred billions of years ago? Or did it rather instantiate presence *then* despite becoming only now an object of my visual perception? If one adopts SVM, these difficulties do not arise. I am not sure whether every part of a perceptual object is, itself, a perceptual object, but I have no doubt that something can have mental states without all of its parts having mental states (the tip of my nose does not have any mental states). Clearly, the same should be true of MENTAL states. I am not sure what to say about events that occur at one time and are perceived at a different time, but I am pretty sure that no pain can occur at one time and be experienced by its subject at a different time (for a pain to occur *just is* for its subject to experience it). Clearly, the point carries over to PAIN and other MENTAL states.

Another reason I have for preferring SVM to EP has to do with what Hare does *not* say about presence. For one thing, he offers no explicit analysis of the notion of "presence" in terms of other, more familiar notions. For another, he doesn't posit any metaphysical or causal connection between presence and other, more familiar properties. So besides being conceptually primitive, presence is also metaphysically insulated. The combination of both features seems to me to be undesirable. Subjectivists are bound to treat the notion of "point of view" (or, alternatively, some other cognate notion, like the notion of something being "otherpersonally" the case) as primitive. I take it that it would be nice if they did not have to

17 The question is McDaniel's: "If x is present, are each of x's parts present?" (McDaniel 2012, 406).

do the same thing with the notions of the particular subjective properties they posit. As to the properties themselves, it would be nice if they could take them to ground or be grounded by other, more familiar properties, instead of reducing them to metaphysical danglers.

Defenders of SVM are not committed to treating MENTAL states in the way Hare treats presence. On the contrary, what I suggested in the last section is that they should connect mental states and MENTAL states by saying that the metaphysical nature (or essence) of any ordinary mental state consists in its being *either* a MENTAL state *or* an otherpersonal MENTAL state. We can now appreciate two immediate advantages of that strategy. The first is that, far from treating MENTAL states as metaphysical danglers, SVM can give them center stage: MENTAL states can be seen as the subjective foundations or 'grounds' of ordinary mental states. The second is that, instead of presenting the notion of a MENTAL state as primitive, a defender of SVM can use various instances of the schema above to elucidate and clarify that notion.

Setting aside my reasons for preferring SVM to EP, I should say that a major difference between Hare and me concerns our respective motivations for endorsing the view we endorse. Hare's motivations are of two sorts. On the one hand, he is interested in vindicating a position that McDaniel (2012) labels 'Rational Egocentric Hedonism', according to which “each one of us should pay special attention to our own well-being” (Hare 2009, 57). One central message of Hare's book is that, unless we adopt EP, we cannot make rational sense of our egocentric-hedonist moral inclinations to favour conditions in which we (as opposed to anyone else) are better off. On the other hand, Hare thinks that EP offers a neat solution to certain puzzles concerning personal identity over time, allowing us to explain the nature of future-directed concerns, desires and expectations in unusual cases involving fission, fusion, brain swapping etc. Overall, then, one could say that Hare's driving motivations for Egocentric Presentism are *practical* in kind: they stem from intuitions concerning what each of us should do, care or be concerned about – in a word, intuitions about what *matters* to each

of us.

My motivations for endorsing SVM are of a different sort. As I said in § 2, I came to embrace SVM because of certain intuitions I have about myself and my place in reality, but if someone pressed me to provide reasons that are more strictly philosophical, the reasons I would cite would not be practical, but *epistemological*. To anticipate a theme that will be central in the second part of this thesis, I think that our knowledge of mental facts exhibit certain peculiar features. And I think one cannot easily explain why our knowledge of mental facts exhibits these features unless something like SVM is true.

CHAPTER 2:

Defending Subjectivism

In the first chapter, I articulated Subjectivism and the Subjectivist View of the Mental (SVM). The aim of this chapter is to defend the tenability of these positions. Upon hearing the thesis that there are subjective facts or the claim that only one subject has MENTAL states, several objections naturally come to mind. In what follows, I will present some of these objections and explain how I think they should be answered. The objections will be organized around four themes: the charge of solipsism (§ 1), the complaint that Subjectivism and SVM are implausibly 'inegalitarian' positions (§ 2), the problem of accounting for intersubjective phenomena like communication and disagreement (§ 3) and the relationship between Subjectivism and Physicalism (§ 4).

Given that it deals mainly with objections, the spirit of the chapter is more defensive than constructive. But it is certainly not *purely* defensive. More often than not, replying to a hostile argument satisfactorily requires more than just blocking the argument: one has to say something positive about the subject matter of the argument, and this may involve taking on commitments that, though not logically built into one's position, flow naturally from it. This will happen at various points in this chapter: what I will be offering is not only a defence of Subjectivism and SVM from some objections, but also an exploration of some views and positions that flow naturally from Subjectivism and SVM, even if they are not logically entailed by them.

1. Subjectivism and Solipsism

In his introduction to Hare's (2009) monograph, Johnston makes the following observation about Hare's view:

> *On Myself, and Other, Less Important Subjects* offers the philosophically most sophisticated form of solipsism (from *solus ipse* – oneself alone) that I have encountered. This is not the crude, almost universally rejected, solipsism that denies the existence of other minds, [but] a deeper and more disturbing solipsism to the effect that the experiences of others are just not present. (Johnston 2009, xi)

Solipsism is usually regarded as a far-fetched view. As Johnston points out, virtually all philosophers reject it. Many have described it as incoherent. So I take it that, in saying that Egocentric Presentism (EP) is a "sophisticated", but also "deep" and "disturbing" form of solipsim, Johnston is implicitly inviting his readers to resist the views and arguments that Hare puts forward in his book. I also take it that, whatever reasons Johnston has for regarding EP as a form of solipsism are equally reasons for regarding SVM as such. After all, the main difference between Hare's view and mine is that for Hare subjective facts involve the instantiation of a primitive property of *presence*, whereas for me subjective facts involve the instantiation of various kinds of MENTAL properties. It is difficult to see how the differences between Hare's presence and my MENTAL properties – however important – could make for a difference between a view that is solipsistic and one that is not. So I think I should begin this chapter by explaining why I think that, pace Johnston, subjectivists like me should not be described as solipsists in disguise.

Different theses go under the name of "solipsism". In *Other Minds* (2001), Avramides characterizes solipsism in the following way:

> Solipsism in its most radical form is a position that rejects all but what is present to the consciousness of an individual. One of the things ruled out by solipsism is the existence of other minds (the other things that are ruled out are other, non-mental, things and other, non-present times) (Avramides 2001, 3)

As Avramides thinks of it, solipsism rules out the existence of all sorts of non-mental things, including mountains and rivers, plants and animals, stars and planets (or, at least, mountains, rivers, plants, animals, stars and planets that are not 'present to anyone's consciousness'). Now I think it should be clear that Subjectivism and SVM do not imply anything so radical and idealistic. Nor is this what Johnston has in mind when he speaks of "solipsism". On his view, Hare's theory is solipsistic in spirit not because it denies the existence of other minds (it doesn't) but because it says "that the experiences of others are just not present". A criticism very similar to this could certainly be raised against SVM, too: SVM is solipsistic in spirit, not because it rejects all but what is present to Giovanni's consciousness (it doesn't), but because it denies that the mental states of subjects other than Giovanni are MENTAL states.

The question is whether this criticism is justified; and I do not think it is. It is true that, according to SVM, only Giovanni has BELIEFS, HOPES, DESIRES, FEELINGS and EXPERIENCES. But that is just one side of the coin. The other side is that, according to SVM, only subjects other than Giovanni have otherpersonal BELIEFS, otherpersonal HOPES, otherpersonal DESIRES and otherpersonal EXPERIENCES. This being so, it is entirely unclear why SVM should be seen as more solipsistic than a view on which there are neither MENTAL states nor otherpersonal MENTAL states. Perhaps, my critics see my view of the mental as 'solipsistic' because it posits *more* properties than theirs and, more specifically, because it says that I am the only bearer of some those properties. But I could just as well describe *their* view of the mental as 'solipsistic' because it posits *less* properties than mine and, more specifically, it says that the subjects I

call “others” do *not* in fact have many of the properties I take them to have. Obviously, I do not pretend to have an unbiased take on the dialectic here. I just think that, instead of using the charge of solipsism as a cudgel, my opponent and I would do much better to admit that the categories of 'solipsist' and 'non-solipsist' do not cut much ice in this context. The distinction here is not between solipsists and non-solipsists. At most, it is between egalitarian views – views according to which all subjects are on a par – and inegalitarian views – views according to which it is *not* the case that all subjects are on a par. The latter distinction and its significance for assessing SVM are matters to which I shall return below (§ 2).

At this point, an objector might suggest that the real reason why Hare's view is a deep and disturbing form of solipsism is not the one Johnston mentions. Maybe the problem does not have to do with any particular claim Hare makes about other subjects and their experiences, but rather with a certain thesis concerning the *points of view* of other subjects – the thesis that such points of view do not exist. This is a thesis that Hare explicitly endorses (“S-worlds” is Hare's name for what I call “points of view” and “a-relation” is his name for a relation of accessibility holding among S-worlds):

> I do not believe that there exists a network of a-related *S*-worlds, and the one that I inhabit is a-related to some other one in which [for instance] Henry Kissinger’s experiences are present. I believe that all that exists is an S-world, S_{ME}, in which the experiences of one person, the person I call “me,” are present. That’s it. (Hare 2009, 27)

Let me call this view “Solipsism about points of view”, to distinguish it from the ordinary form of solipsism about subjects, which denies the existence of other minds:

(*Solipsism about points of view*) There are no points of view other than the firstpersonal one.

Precisely because I take Hare to be a subjectivist, I am surprised that he endorses *Solipsism about points of view.* Recall what Subjectivism says: some propositions are true simpliciter without being true from all points of view. If there were no points of view other than the firstpersonal one – as *Solipsism about points of view* says – *nothing* could be true simpliciter without being true from all points of view. To put it otherwise, if there were no points of view other than the firstpersonal one, reality would not be subjectively the way it is – it would be the way it is from *every* point of view. That is exactly the opposite of what subjectivists hold. So I personally reject *Solipsism about points of view.* And I think that Hare had better reject it, too.

The point is, I think, worth emphasizing. The distinctive feature of solipsism is generally taken to be "the inability to make sense of the idea of real minds [or points of view] other than one's own" (Nagel 1986, 20). Since the very formulation of Subjectivism makes reference to the existence of points of view other than the firstpersonal one (or, what amounts to the same thing, to the difference between what is firstpersonally the case and what is otherpersonally the case), SVM should be seen as the exact *opposite* of solipsism: the recognition that the way things are *for oneself* does not coincide with the way things are *for others* lies at the very core of SVM and represents one of the key motivations for embracing it. As a defender of SVM, then, I want to say – and am perfectly entitled to say – what Nagel says in this passage:

> Not being a solipsist, I do not believe that the point of view from which I see the world is *the* perspective on reality. Mine is only one of the many points of view from which the world is seen. (Nagel 1986, 57)

Of course, this does not mean that SVM puts all points of view on the same level. As any version of Subjectivism, SVM says that, of all the points of view there are, only the

firstpersonal point of view is such that any proposition true relative to it is also true simpliciter. But just as the mere fact that I draw a certain distinction between myself and everyone else does not make me a solipsist about subjects, the fact that I draw a certain distinction between firstpersonal truth and otherpersonal truth does not make me a solipsist about points of view. At most it makes my view somewhat *inegalitarian* in its treatment of points of view. So, putting the accusations of solipsism to one side, let me turn to questions having to do with egalitarianism and its denial.[18]

2. Subjectivism and Inegalitarianism[19]

I've been arguing that neither Subjectivism nor SVM imply any form of solipsism. But I've granted that both Subjectivism and SVM may be accused of being somewhat *inegalitarian* views. SVM is inegalitarian in its treatment of subjects: it says that one subject is special, because his mental states are MENTAL states. As to Subjectivism – the general view that some propositions are true simpliciter without being true from all points of view – it is inegalitarian in its treatment of points of view, in the following sense:

1) A theory is not egalitarian in its treatment of points of view if there is some point of view or another that it discriminates against.
2) A theory discriminates against a point of view s if, for some proposition p, the theory

18 The solipsistic theses I've been looking at are all metaphysical theses connected with the idea that one is the only subject in the world. There are also epistemological versions of solipsism, saying that it is impossible to know anything about other people's minds. For a discussion of the relationship between epistemological solipsism and the Subjectivist View of the Mental, see Chapter 3, § 4.

19 This section draws on ideas already presented in Merlo (2013).

fails to affirm that p is true even if p is true from the point of view s in question.

3) Any theory on which some propositions are true without being true from all points of view is a theory that fails to affirm the truth of certain propositions even if there are points of view from which those propositions are true.
4) Any theory that accepts Subjectivism discriminates against some point of view or another (by (2), (3) and the definition of Subjectivism)
5) Any theory that accepts Subjectivism is not egalitarian in its treatment of points of view.

   (by (1) and (4)).

Of these two forms of inegalitarianism – inegalitarianism about subjects and inegalitarianism about points of view – the second seems to me to be the one that subjectivists should focus their attention on. This is not to say that inegalitarianism about subjects is completely unproblematic, but if there is something to the intuitions I discussed in Chapter 1, when I explained my motivations for embracing SVM, egalitarianism about subjects cannot be made out to be unproblematic either. What's more, the species of inegalitarianism about subject that SVM entails is not *objective* inegalitarianism: Giovanni is said to be special, but only subjectively so, in the sense that there are points of view from which he is not special. If even *after* we factor in this observation, the feeling persists that SVM is disturbingly 'inegalitarian' in its treatment of subjects, the problem must have to do with the fact that the points of view from which Giovanni is not special are not treated 'on a par' with the point of view from which Giovanni is special – that they are somehow discriminated against. This suggests that, insofar as they want to shake off the reputation of being “inegalitarian”, it is inegalitarianism about points of view that advocates of SVM should be primarily concerned with.

The question then is: *should* a friend of SVM try to shake off the reputation of having an "inegalitarian" attitude towards points of view (and, therefore, towards subjects)? In a sense, yes, and in another, no. Admittedly, there isn't much leeway to resist the argument above: premise (1) does little more than articulate what inegalitarianism about points of view amounts to, premise (3) is unobjectionable[20] and it can hardly be denied that there's a way of reading the expression "discriminating against a point of view" on which premise (2) makes perfect sense. This is just to say that, if they are to be honest and upfront about their credo, subjectivists should not be afraid of admitting that *a certain kind* of inegalitarianism about points of view is built right into their position. Again, some will cry scandal and call this "madness". I don't think there is much that subjectivists can (or should) do in reply, except keep their heads cool, stick to their guns and wait for more polite reactions.

But then again, this need not be the end of the story. After all, there *is* some intuitive pressure to say that all points of view are 'equally real' and it would be nice if subjectivists could show that they are not completely insensitive to that pressure. One thing they could do – and perhaps the only thing they could do – is remind their critics that a theory can be inegalitarian in one respect and perfectly egalitarian in another. More precisely, they could argue that – when it comes to vindicating the intuitive sense in which all points of view are 'equally real' – failing to affirm the truth of a certain proposition need *not* be a way of discriminating against the points of view from which that proposition is true. In other words, they could argue that, when "discriminating against a point of view" is read in the particular way which speaks to our egalitarian feelings, (2) can be plausibly rejected. In the remainder of this section, I shall present a few ways in which this could be done.

20 Take any proposition p that, according to the theory, is true without being true from all points of view: the proposition *that p is not true* is true from some point of view, but of course the theory will not affirm its truth.

The minimalist strategy

I want to start with the idea of rejecting (2) on the basis of:

(*) A theory only discriminates against a point of view S if, for some proposition p, *there is no point of view from which the theory affirms that p* and yet p is true from the point of view S in question.

The gist of the idea is simple. According to Subjectivism, truths varies from one point of view to another. This suggests that subjectivists should be especially interested in theories whose content *also* varies from one point of view to another. For example, take SVM. As I formulated it at the end of the last chapter, SVM affirms that Giovanni is special and you are not. That is what the theory says from the firstpersonal point of view, which is also what the theory says simpliciter. But what does the theory say *otherpersonally*? More specifically, what does the theory assert *from your point of view*? Intuitively, it should say that you are special and I am not. *That* is what is true from your point of view. *That* is what I am trying to convince you of when I am trying to convince you to accept SVM. There is no reason why a theory that says certain things from one point of view could not be stipulated to say different things from another point of view, just as there is no reason why a theory that says certain things at one time could not be stipulated to say different things at a different. Take Presentism, the philosophical view that everything exists presently. The tacit understanding that, at every future time t, Presentism will say about t what it now says about the present time is part and parcel of our grasp of this view. Mutatis mutandis, the same applies to SVM: from the point of view of any subject x, SVM says about x what, from the firstpersonal point of view, it says about Giovanni.[21] Now, the crucial thought is that, when the content of a theory

21 The idea is that SVM is associated not with a set of propositions, but rather with a function from points of

varies in this way, the theory need not state *from every point of view* everything that is true *from some point of view or another*: it will be egalitarian enough if it ensures that every point of view gets, *from some point of view or another*, the share of fame it deserves. This is why, by the lights of a subjectivist, (2) can be plausibly rejected on the basis of (*).

This response succeeds in blocking the argument and is pleasingly 'minimalist', in the sense that it requires no fundamental revision of the basics of Subjectivism. Yet there is something unconvincing about it.

First of all, the response draws on the fact that points of view other than mine (e.g. *your* point of view) get the share of fame they deserve *from points of view other than mine* (e.g. from *your* point of view). But it might be complained that what SVM says *from other points of view* is quite irrelevant here. The problem is with what the theory *does* say and what it does say is pretty discriminatory against any point of view other than mine. Considered in the light of this simple objection, (*) doesn't carry much conviction.

Another problem with the 'minimalist' response is the following: to the extent that it succeeds in the subjective case, one would expect it to have some bite in the temporal and in the modal case, too. But can a theory of time describe itself as 'egalitarian' in its treatment of the time series just because, *sooner or later*, every time will have been described by that theory as the present time? I'm not sure. And can a theory of modality describe itself as 'egalitarian' in its treatment of possible worlds merely because every possible world *could* be described by that theory as the actual world under certain metaphysically possible circumstances? I wouldn't think so. Now, to be sure the mere observation that the 'minimalist' response doesn't carry over to the temporal and modal case doesn't mean that it fails in the subjective one. But it certainly puts some pressure on its proponent to explain *why* it doesn't

---

view to sets of propositions. One may ask whether different subjects *disagree* with one another when they all embrace the theory, given that what the theory says changes from one point of view to another. The question will be addressed below (§ 3).

carry over – and this just means that it puts some pressure on its proponent to go beyond its pleasing 'minimalism'.

Thirdly, and relatedly, the minimalist response makes egalitarianism about points of view implausibly cheap, if not inescapable. Call 'complete' a theory that, from each and every point of view, states everything that is true from that point of view. By (*), any 'complete' theory will be perfectly egalitarian in its treatment of points of view. This seems an undesirable result: whether or not all points of view are 'equally real' should be a substantive metaphysical question and it should in principle be possible for a subjectivist to formulate a 'complete' subjectivist theory without answering that question in the positive. This is another reason not to accept (*).

I take these difficulties to show that, insofar as they want to reconcile their view with the egalitarian intuition that all points of view are 'equally real', subjectivists must do more than just stress the subjective nature of their theories. But what more could they do?

The Fragmentalist strategy

A more robust form of egalitarianism could perhaps be achieved by endorsing a version of Kit Fine's (2005) *Fragmentalism*. There are two ingredients to *Fragmentalism*. The first is a robust distinction between reality itself – which may be thought of as the totality of what is the case – and *metaphysical reality* – which may be thought of as the totality of what is *really* the case. Fine says that "whatever is really the case (belongs to metaphysical reality) may, with some plausibility, be taken to be the case (belong to mere reality). But the converse will not in general hold; and so […] I might accept that I am sitting and even accept that it is a fact that I am sitting,but not accept that this fact is constitutive of how things really are" (Fine

2005, 267). The second ingredient is the idea that metaphysical reality (or "meta-reality", as I proposed to call it in the last chapter) is genuinely incoherent: for some propositions p and q incompatible with one another, it is really the case that p and it is really the case that q. This leads to a radically new view, on which "it is taken to lie in the character of [meta-reality] that certain apparently contradictory aspects of it cannot be explained away" (Fine 2005, 281).

Using Fine's distinction between what is the case and what is really the case, a subjectivist who endorses *Fragmentalism* can propose to reject (2) on the basis of :

(**) A theory only discriminates against a point of view S if, for some proposition p, the theory fails to affirm that *it is really the case that* p even if, from the point of view S in question, *it is really the case that* p.

She can then make sure that no point of view is 'discriminated against' by endorsing:

(*Stability*) For any proposition p, if otherpersonally it is really the case that p, then it is really the case that p.

Of course, what is really the case from one point of view need not be compatible with what is really the case from another point of view – which is why, in a subjectivist framework, a principle like *Stability* will have the distinctive consequence of making reality 'incoherent'.

Fragmentalism is a fascinating view, but I don't think it can form the basis of a successful reconciliation of Subjectivism with the egalitarian intuition that all points of view are 'equally real'. First of all, notice that Fine's notion of what is really the case is factive: whatever is really the case is the case (as Fine says elsewhere the concept of meta-reality "enables us to distinguish, *within the sphere of what is the case*, between what is really the

case and what is only apparently the case" (Fine 2001, 3; my emphasis)). This means that any incoherence will inevitably spread from meta-reality to reality: two propositions p and q incompatible with one another can both turn out to be true, if both of them are really the case. It is difficult to see how the fragmentalist can possibly avoid the spectre of true contradictions (the fragmentalist can, of course, insist that no true contradiction is *really* true, but this seems to me to be little consolation).[22] Secondly, notice that, given *Stability*, the fragmentalist won't admit any kind of subjective variation in what is really the case. Now, this would still be compatible with subjective variation in reality (i.e. in the totality what is the case), if the notion of what is really the case were not factive. But since it is factive, there's a genuine risk that stabiliy, too, will spread from meta-reality to the reality, with the result that our fragmentalist version of Subjectivism will start to look dangerously similar to a view on which Subjectivism is false and reality is just objectively contradictory.

The grounding strategy

Not everything is lost. As it stands, the fragmentalist strategy is unsuccessful, but its starting point seems to me to be sound: subjectivists need something like Fine's distinction between what is the case and what is really the case, if they want to be honest and upfront about the inegalitarian nature of their view about reality and, *at the same time*, say that there is a legitimate sense in which all points of view are 'equally real'. In other words, if (2) is to be rejected, the kind of principle on the basis of which subjectivists should reject it.

Moreover, even if the fragmentalist strategy is unsuccessful, its failure is instructive.

22 Fine points out "although there is a sense in which the fragmentalist takes reality to be contradictory, […] it will not be correct for me to assert both that I am sitting and that I am standing" (2005, 282). But it is not clear to me whether the point is supposed to concern *assertibility* rather than *truth*.

The lessons to be learned are essentially two. The first is that any egalitarian version of Subjectivism had better respect *Stability*, the principle that whatever is really the case from any point of view *is* really the case. This much fragmentalist subjectivists were right about: unless meta-reality is 'stable' (and, therefore, incoherent), we cannot even begin to make sense of the idea that, when it comes to how real they are, other points of view are perfectly 'on a par' with the firstpersonal one.

The second thought is that stability and incoherence will inevitably spread from meta-reality to reality, unless we deny something that Fine's fragmentalist accepts, namely:

(*Factivity of meta-reality*) If it is really the case that p, then p.[23]

Rejecting this principle is certainly a controversial move, but one that the egalitarian subjectivist should not, I think, try to avoid. For if one wants to include the otherpersonal *as such* in meta-reality – if one wants to say that something may be really the case while being only *otherpersonally* the case – one shouldn't be afraid of saying that meta-reality extends *beyond* the limits of reality (i.e. the totality of what's the case): certain things that are not the case are, nonetheless, really the case.

---

23 In discussing Fine's fragmentalist view of time, Rosenkranz and Correia point out that "one reason to be unhappy with fragmentalism is that it is incompatible with a principle which has a great pre-theoretical plausibility, namely:

(*Truth*) If at a given time t, f constitutes reality and is the fact that p, then it is true at t that p" (Rosenkranz and Correia 2012, 311).

But, as we've seen, Fine affirms that "whatever is really the case (belongs to metaphysical reality) may, with some plausibility, be taken to be the case" (Fine 2005, 267). So my take on Fine's Fragmentalism is that it actually incorporates a commitment to *Truth*. For a discussion of these issues, see also Rosenkranz and Correia (2011).

Of course, the two lessons need to be supplemented with a *pars construens*: after rejecting *Factivity of meta-reality*, the egalitarian subjectivist must provide us with some independent gloss on the notion of meta-reality she has in mind, otherwise her acceptance of *Stability* will have the air of an unprincipled stipulation. What could the alternative and independent gloss be? Here one promising line starts with the idea that whenever what a certain proposition p says *metaphysically explains* – or 'grounds' – what another proposition q says, there's a rather natural sense in which what p says is really the case, for it is by reference to it that we understand how things most fundamentally are with respect to q.[24] This provides intuitive support for:

(*Grounding*) For every proposition p, it is really the case that p if and only if, for some proposition q, p grounds q.

With *Grounding* in place, it's fairly clear what the egalitarian subjectivist needs to do to substantiate her position: she needs to show that there are distinctively metaphysical or 'grounding' explanations featuring propositions that are only otherpersonally true as explanans.

How could this be? What kind of propositions could possibly be grounded or metaphysically explained by *false* propositions (albeit ones that are otherpersonally true)? The answer is, I think, pretty obvious: true propositions about what is *otherpersonally* the case. Take the (true) proposition *that, otherpersonally, Paul is special* (let Paul be a subject distinct from Giovanni). This proposition may be true, but it is natural to think that truths of this sort stand in need of metaphysical explanation.[25] Now, suppose that, otherpersonally, it is really

24 The idea should be familiar. See Correia and Schnieder (2012).

25 An analogous worry arises for true propositions about what was or will be the case. This is sometimes put by saying that the truths in question are 'hypothetical' rather than 'categorical'. The distinction is due to Sider

the case that Paul is special. Given *Grounding*, this is tantamount to saying that, otherpersonally, the proposition *that Paul is special* grounds or metaphysically explains other propositions. But then, why should we not use precisely that proposition to explain why, otherpersonally, Paul is special? We could say that *otherpersonally, Paul is special* because, from some other point of view, *Paul is special*. Here a firstpersonal truth about what is otherpersonally the case (namely, the true proposition *that, otherpersonally, Paul is special*) is explained in terms of a proposition that is only otherpersonally true (namely, the proposition *that Paul is special*). In other terms, a fact about what is otherpersonally the case is explained, not by appeal to some other fact, but rather in terms of what is otherpersonally the case.[26]

The idea can be formalized using an idiom that, alongside 'OTHER', contains two more sentential operators. One is a non-factive operator 'R' meaning 'It is really the case that...'. The other is a two-place (but potentially many-place) operator '*OTHER-BECAUSE*', such that statements of ground featuring this operator (e.g. "p *OTHER-BECAUSE* q") can be true despite the sentence (or sentences) stating the ground being only otherpersonally true. The general principle that the egalitarian subjectivist needs to endorse is:

(*Intersubjective Ground*) For any proposition p, [OTHER R(p) → (OTHER (p)) *OTHER-BECAUSE* (p))]

This says that if, otherpersonally, it is really the case that p, it it by reference to p that one

(2001, 35-42).

26 Similar moves can be made in the temporal case. Present truth about the past (e.g. the proposition *that Caesar crossed the Rubicon*) can be explained in terms of propositions that used to be true but are no longer such (e.g. the proposition *that Caesar is crossing the Rubicon*). Sanson and Caplan (2010) seem to me to propose something along these lines. Symmetrically, present facts as to what will be really the case can be explained by appeal to how things *will be* – this is, in effect, the strategy adopted by what Rosenkranz (2012) calls the 'Ockhamist'.

should explain OTHER p.[27] Given *Grounding*, *Intersubjective Ground* entails precisely:

(*Stability**) For every proposition p, OTHER R(p) → R(p)

which is what the egalitarian subjectivist wanted: whatever is really the case from other points of view *is* really the case.

Notice that *Intersubjective Ground* is a substantive metaphysical principle: it affirms that the metaphysical explanation of at least certain kind of facts cannot be found *within* the firstpersonal point of view, but lies, as it were, *outside* reality (i.e. the totality of what is the case), in the non-factive realm of what's otherpersonally the case. The principle, if true, may well be necessarily true, but whether it is true is an issue that subjectivists of different strands will disagree on. And this is just as it should be: egalitarianism should not come on the cheap, let alone be inescapable.

There are two other reasons for thinking that tying the truth of *Intersubjective Grounding* to the egalitarian intuition that all points of view are 'equally real' is the right thing to do. First, one might take *Intersubjective Grounding* to say that no point of view is like a world apart: to make metaphysical sense of everything that is the case from my point of view I need to look *also* at what happens from your point of view and to make metaphysical sense of everything that is the case from your perspective, you have to look *also* at what's the case from my perspective. Even before principles like *Grounding* and *Stability* are brought in, this strikes me as a very 'egalitarian' theme: your and my points of view are metaphysically connected in such a way that, to the extent that one of them participates in meta-reality, *both*

27 Notice that *Intersubjective Ground* does not imply that *every* truth about what is otherpersonally the case is grounded in what is otherpersonally the case. Intuitively, at least some propositions concerning what is otherpersonally the case are grounded in what is the case (for example, the proposition *that, otherpersonally, it is otherpersonally the case that Giovanni is special*).

of them must do so.

Second, there seems to me to be a rather sharp contrast between *Intersubjective Grounding* and its modal analogue:

(*Inter-world Grounding*) For any proposition p, [POSSIBLY R(p) → (POSSIBLY (p))

POSSIBLY-BECAUSE (p))]

Inter-world Grounding says (roughly) that the fact *that the atom could decay* should be metaphysically explained by appeal to the proposition, true only in some other possible world, *that the atom decays* ('*POSSIBLY-BECAUSE*' is supposed to be an operator such that statements of ground featuring it can be true even if the sentence that states the ground is only possibly true). Now, explanations of this kind don't sound very natural to us: there's no natural account of the actual in terms of the merely possible (at best, the explanation goes the other way around). This observation – the lack of metaphysical interconnectedness among different possible worlds – may contribute to explain something that the 'minimalist' response considered above failed to explain, namely the asymmetry between the subjective case (where egalitarianism has some intuitive pull) and the modal case (where egalitarianism has little or no pull).[28]

Of course all this is not to say that the strategy I've sketched has no costs. The idea that falsehoods can do essential explanatory work is certainly not new, but it should be conceded that there is something revisionary in applying it to the case of metaphysical or 'grounding' explanations.[29] My tentative suggestion here is that this revisionariness might be

28 I take the temporal case to be intermediate between the two, but closer to the subjective case. See Merlo (2013).

29 For example, it has been argued that mathematical statements do essential explanatory work despite being false (Leng 2010).

the price we need to pay to reconcile Subjectivism with the egalitarian thought that all points of view are, in some sense, 'equally real'.

3. Subjectivism and Intersubjectivity

The fact that a principle called "Intersubjective Grounding" may be thought to bring some grist to the subjectivist mill is somewhat ironic. For one would have thought that, far from being a source of help, the notion of intersubjectivity would pose a *challenge* to Subjectivism. After all, there is a sense in which, for a subjectivist, whatever happens (or exists or is the case) happens (or exists or is the case) *within* a particular point of view: the firstpersonal one. On the other hand, the range of phenomena that go under the label of "intersubjectivity" are supposed to concern what happens (or exists or is the case) *in between* one point of view and another.[30] The use of operators like *OTHER-BECAUSE* is an example of how subjectivists might deal with a particular instance of this challenge: the operator allows one to state explanatory truths that, while holding (as any other truth) from the firstpersonal point of view, are supposed to connect first- and otherpersonal truths.[31]

But the challenge arises in other areas, too. How is a subjectivist going to account for the possibility of *communication* across different points of view? What kind of account of intersubjective *disagreement* can a subjectivist provide? Can we even make *rational* sense of each other's beliefs if Subjectivism is true and the beliefs we form are formed from different

30 For analogous reasons, cross-temporal relations are generally held to pose a challenge to Temporaryism in general and Presentism in particular (see Sider (2001), Crisp (2005) and De Clercq (2006)).

31 There's an interesting analogy here with the use of so-called 'span operators' in a temporaryist framework. See Brogaard (2007).

points of view? These questions are closely interrelated and in this section I will try to make some suggestions as to how subjectivists should answer each of them. I will start with the last one, which may be construed as a question about the relation between Subjectivism and the chief principle we use in interpreting one another, the principle of Charity.

Charity

SVM says that I am special and that my mental states are MENTAL STATES. But I bet that, after rehearsing the intuitions of Chapter 1, you convinced yourself that *you* are special and that *your* mental states are MENTAL STATES. Now let it be granted that (as I suggested in the last section) one can speak of *one* theory here that you and I are both believing – a theory that says certain things from my point of view and different things from your point of view. Also, forget for a moment the question whether you and I can be said to *agree* with one another when we both espouse this theory (the question will be taken up below). The point remains that, from my point of view, some of the propositions you hold are blatantly false: you are not special and your beliefs are not BELIEFS. Symmetrically, from your point of view, some of the propositions I believe are blatantly false, because from your point of view, I am not special and my beliefs are not BELIEFS. But now consider the chief principle we use in interpreting other rational subjects:

(*Charity*) If x is a rational subject, then it is not the case that x has any inexplicably wrong beliefs.[32]

32 I formulate *Charity* as a *truth* concerning rational subjects. According to other formulations, the principle of charity is not so much a truth, but a regulative principle or a heuristic constraint for interpreting other subjects. The principle is prominent in the work of Davidson (1973; 1983).

It seems to follow from *Charity* that, unless I can explain why you came to hold the wrong beliefs you hold (and unless you can do the same from your point of view), we have to regard each other as irrational subjects. The problem is: how can *I* possibly explain the fact that you believe to be special when, from my point of view, it's so clear that *I* am special and you're not? And how can *you* possibly explain the fact that I believe to be special when, from your point of view, it's so clear that *you* are special and I am not? We are caught between a rock and a hard place: explaining something that we don't have the resources to explain or denying each other the status of rational subjects. What's the way out?

To answer this question, we need to carefully reconstruct the argument from Charity. A first version of the argument aims at establishing that, given any proposition p that is true from my point of view but false from yours (like the proposition *that Giovanni is special*), if I believe p then from your point of view I am not rational:

(1) I believe p and, from your point of view, p is false;
(2) From your point of view, I believe a false proposition;
(3) From your point of view, I hold an incorrect belief;
(4) From your point of view, either I am irrational or there is an explanation of why I believe p;
(5) But from your point of view, there is no explanation of why I believe p.
(6) So from your point of view, I am irrational.

A parallel argument can be made that, given any proposition q that is true from your point of view but false from mine (like the proposition *that you are special*), if from your point of view you believe q then from my point of view you are not rational:

(1') From your point of view, you believe q, which is false from my point of view;

(2') You believe a false proposition;

(3') You hold an incorrect belief;

(4') Either you are irrational or there is an explanation of why you believe q;

(5') But there is no explanation of why you believe q.

(6') So you are irrational.

It seems to me that the key moves are two: the move from (1) to (2) (and from (1') to (2')) and the move from (2) to (3) (and from (2') to (3')).[33] The first move takes us from the claim that I believe p to the claim that I do so from your point of view (and, symmetrically, from the claim that you believe q from your point of view to the claim that you believe q). This requires a principle that might be called:

(*Objectivity of belief*) If x believes a proposition p, then it is true from all points of view that x believes p.

The second move takes us from the claim that I (or you) believe a false proposition to the claim that I (or you) hold an incorrect belief. This follows from:

(*Correctness*) A belief in a proposition p is correct if and only if p is true.

I don't think that subjectivists should reject either of these two principles *sic et simpliciter*.[34]

33 (5) and (5') are also something that, as a subjectivist, one would like to question. But I think that the best way of questioning (5) and (5') is by looking at (1) and (2) and (1') and (2').

34 In this sense, I disagree with Hare (2009, 52-55), who suggests that denying (what I call) *Correctness* or

What they should do is use the distinction between the ordinary notion of belief and the subjectivist notion of BELIEF to formulate different versions of these principles. The arguments above should then be faulted on grounds of equivocation: when run in terms of one notion, they fail at step (2) (and (2')); when run in terms of the other notion, they fail at step (3) (and (3')).

For example, suppose one tries to run the arguments above using the subjectivist notion of BELIEF. Then what is needed are not *Objectivity of belief* and *Correctness*, but their capital-letter analogues, namely:

(*Objectivity of BELIEF*) If x BELIEVES a proposition p, then it is true from all points of view that x BELIEVES p.

(*CORRECTNESS*) A BELIEF in a proposition p is correct if and only if p is true.

And what should a friend of Subjectivism say about these principles? It seems to me that she should grant that *CORRECTNESS* is true (indeed, objectively true). Even better: she should say that part of what makes BELIEF the most fundamental kind of belief (a kind whose notion deserves to be written in capital letters) is the fact that it is governed by a simple principle like *CORRECTNESS*: a mental state is a BELIEF only if it is sensitive to truth in the straightforward way described by *CORRECTNESS*. But precisely because BELIEF is sensitive to truth and truth (in a subjectivist framework) can be subjective rather than objective, a good friend of Subjectivism should deny *Objectivity of BELIEF*: for example, I BELIEVE certain things but I fail to do so from your point of view and, symmetrically, it is not the case that you BELIEVE certain

Objectivity of belief are two completely alternative moves open to the subjectivist.

things, although it may be the case that you do so from your point of view.[35] So when the arguments above are stated in terms of BELIEF, they fail at step (2) (and (2')): I BELIEVE myself to be special, but it is not otherpersonally the case that I do so, so you cannot take me to have incorrect BELIEFS (symmetrically, you don't BELIEVE yourself to be special – it is only otherpersonally the case that you do so – and therefore I cannot take you to have incorrect BELIEFS).

On the other hand, suppose that someone tries to run the argument in terms of the ordinary notion of belief and, therefore, relies on the original formulations of *Objectivity of belief* and *Correctness*. In such a case, what should be questioned is not the legitimacy of moving from (1) to (2) (or from (1') to (2')). What should be questioned is, rather, the very truth of *Correctness*. The friend of Subjectivism should remind her interlocutor of the general principle that ordinary mental states have a disjunctive nature. In particular:

believe (x) $\leftrightarrow_R$ BELIEVE (x) v OTHER BELIEVE (x)

To believe is to *either* BELIEVE *or* otherpersonally BELIEVE.

She should then say that, since belief has a disjunctive nature, the correctness conditions of a

35 In principle, the failure of *Objectivity of BELIEF* could be said to result from the failure of either of these two theses:

(*Objectivity of BELIEVING*) If there is some proposition that x BELIEVES, then it is true from all points of view that there is some proposition that x BELIEVES.

(*Objectivity of what is BELIEVED*) If x BELIEVES a proposition p, then it is true from all points of view that, if there is some proposition that x believes, p is one such proposition.

But it seems to me obvious that it is *Objectivity of BELIEVING* that should go. It would be rather odd to think that, rather than failing to hold any BELIEFS from your point of view, I actively BELIEVE something different from your point of view than I do from mine.

belief are also disjunctive:

(*Correctness**) A belief is correct if and only if it is a correct BELIEF *or* it is otherpersonally a correct BELIEF.

With *Correctness** in the place of *Correctness* the arguments above can be blocked at step (3) (and (3')). For instance, suppose you believe yourself to be special: that belief is correct even from my point of view because, otherpersonally, it is a correct BELIEF. Symmetrically, my belief that I am special should be regarded as correct even from *your* point of view, despite having a content which is, from your point of view, false.

In the framework of SVM, *CORRECTNESS* and *Correctness** can be shown to entail a rather elegant principle, namely:

(*Relative Correctness*) For any subject x, if x believes a proposition p then x's belief is correct if and only if p is true from x's point of view.

To see how *Relative Correctness* follows from *CORRECTNESS*, *Correctness** and SVM let us reason by cases. First consider Giovanni. According to SVM, Giovanni's beliefs are BELIEFS. So, by *CORRECTNESS*, they are correct if and only if their content is true, which is just to say that they are correct if and only if their content is true from Giovanni's point of view (Giovanni's point of view being the firstpersonal one, i.e. the point of view from which and all and only what is simpliciter is true). Next consider any individual other than Giovanni. According to SVM, the beliefs of that individual are only otherpersonal BELIEFS. So, by *Correctness**, they are correct if and only if otherpersonally they are correct BELIEFS, which is just to say that they are correct if and only if their content is true from that individual's point

of view (that being the *only* point of view from which they are BELIEFS). So the open sentence in *Relative Correctness* is true of Giovanni and of any individual other than Giovanni. So it is true of any subject. So *Relative Correctness* is true.

The basic idea of the response – that of distinguishing neatly between principles that apply to mental states and principles that apply to MENTAL STATES – can also be used to solve other apparent puzzles. For instance, someone might ask how Paul (an individual distinct from Giovanni) can possibly *know* that he is special given that knowledge is factive and, according to SVM, only I am special. The answer at this point should be obvious. One needs to distinguish three thesis. The first concerns KNOWLEDGE and says that:

(*FACTIVITY*) If someone KNOWS a proposition p, then p is true.

Given that only I have KNOWLEDGE, this thesis only applies (non-vacuously) to what I know. The second thesis, that follows from the first given the disjunctive definition of knowledge in terms of KNOWLEDGE, says that:

(*Factivity**) If someone knows a proposition p, then p is either true or otherpersonally true.

And the third thesis, that follows from the first two given SVM, says that:

(*Relative Factivity*) For any subject x, if x knows a proposition p then p is true from x's point of view.

Given *Factivity**, there is no incompatibility between the Subjectivist View of the Mental and the supposition that Paul knows himself to be special. Given *Relative Factivity*, we know what

has to be the case if Paul knows himself to be special: it has to be the case that Paul is special from Paul's point of view.[36]

If this line of response is the right one, there's a moral that the two arguments above suggest. The moral is *not* – as critics of Subjectivism would want – that there's a tension between Subjectivism and the possibility to recognize and be recognized by other subjects as rational: if they play their cards well, subjectivists have all the resources to make rational sense of each other's beliefs. The moral consists, rather, in a conditional thesis: *if* any proposition at all is merely subjectively true (as subjectivists hold) and *if* these propositions can become objects of various propositional attitudes (as seems inevitable), *then* there is good reason to think that at least some propositions about the distribution of propositional attitudes will be subjective, too. Why? Because propositional attitudes represent reality, and they can do so correctly or incorrectly. If reality is subjectively the way it is, a view that treats the distribution of propositional attitudes as objective will end up with too many inexplicably incorrect attitudes. We don't want that. So we have to treat at least some propositional

36 Some might object that *Relative Factivity* constitutes a radical departure from a principle whose truth should be regarded as unquestionable, namely:

(*Factivity*) If someone knows a proposition p, then p is true.

Now, it's not clear to me that the linguistic evidence is incontrovertibly in favour of *Factivity*, rather than *Relative Factivity*. But even granting that it is, there are several things that subjectivists can do to alleviate worries that their take on the ordinary notion of knowledge is radically revisionary. First of all, they can accept that *Factivity* is true when restricted to the class of objective propositions (i.e. the propositions that are true from all points of view if true at all). Secondly, they can point out that, once it is allowed that the world I inhabit is different from the one you inhabit (i.e. that the world is only subjectively the way it is), *Relative Factivity* is pretty much all we need to capture the intuitive idea that knowledge is the kind of state of mind whereby subjects are successfully related to the world they inhabit. Last, but not least, they can notice that, with respect to the task of vindicating *Factivity*, subjectivists are no worse off than non-subjectivist subject-relativists: any subject-relativist who thinks that subject-relative propositions (e.g. the proposition *that chocolate is tasty*) can be known will need to replace *Factivity* with something along the lines of *Relative Factivity*.

attitudes (namely, the attitudes whose correctness is directly sensitive to truth simpliciter) as subjective rather than objective. This is just another way of saying that, if Subjectivism is true, some version of Subjectivism *about the mental* naturally follows.

Disagreement

A related problem for Subjectivism is the problem of disagreement. The problem arises from what is, prima facie, an uncontroversial principle:

(*Disagreement*) For any subjects x and y, x disagrees with y if, for some p, x believes that p and y believes that not p.

If *Disagreement* is true, I disagree with you if I believe myself to be special (as I do) and you believe yourself to be special (as I encourage you to do). Clearly, this is not a result that subjectivists can be happy with. By inviting each and every of my readers to embrace SVM (i.e. by inviting each and every of my readers to believe what SVM says from his or her point of view), I am aiming at creating consensus, not widespread and (presumably) irresolvable disagreement! And yet on what basis could one reject *Disagreement*?

Taking inspiration from what has been said above about BELIEVING and believing, subjectivists could suggest that *Disagreement* should be replaced by:

(*DISAGREEMENT*) For any subjects x and y, x disagrees with y if x BELIEVES that p and y BELIEVES that not p.

And that, for the ordinary notion of belief, disagreement should be defined derivatively:

(*Disagreement**) For any subjects x and y, x disagrees with y if *either* x BELIEVES that p and y BELIEVES that not p *or* it is otherpersonally the case that x BELIEVES that p and y BELIEVES that not p.

It takes little reflection to see that neither *Disagreement** nor *DISAGREEMENT* license the conclusion that there is disagreement between you and me if I believe myself to be special and you believe yourself to be special. This, however, might be seen as a Pyrrhic victory. For notice that *Disagreement** and *DISAGREEMENT* only make room for the possibility of disagreement between subjects that, from some point of view or another, both have BELIEFS. But SVM denies that there are any *two* such subjects (for every subject s, the only point of view from which s has BELIEFS is one from which *no one else* has BELIEFS). In more vivid terms, one can say that, given SVM, *DISAGREEMENT* and *Disagreement** only make room for the possibility of *intrasubjective* disagreement (disagreement of oneself...with oneself!). But clearly there can be disagreement between me and you, i.e. between someone who has BELIEFS and someone who has only otherpersonal BELIEFS. So now we need some principle telling us under what circumstances such a type of disagreement can occur. In other words, we need a good (i.e. plausible and independently motivated) criterion for *intersubjective* disagreement. How to formulate such a criterion?

A good idea is to look at how we model *diachronic disagreement* (i.e. disagreement across time) in a temporaneist framework. Suppose I now believe that Obama is the president of the United States and suppose that, a century from now, my (unlearned) grandson will believe that a century earlier Obama was not the president of the United States. Intuitively, there is some kind of disagreement between my grandson and me, even if we do not hold our

respective beliefs simultaneously (i.e. if it is never the case that they are both *present* beliefs). But one might want to say that the content of my belief is the temporal proposition *that Obama is the president of the United States* and the content of my grandson's belief the temporal proposition *that a century earlier Obama was not the president of the United States* – two propositions that do not contradict one another. To recover the sense in which my grandson and I disagree, one could look at the sempiternal propositions my grandson and I believe. But what if the only propositions about Obama's presidency that my grandson and I believe are temporal? Then, what one needs to do is look not at the propositions that the two beliefs take as their objects, but rather at what might be called the *atemporal contents* of the two beliefs, where the atemporal content of a belief is defined as follows:

For any belief b in a proposition p which is held at a time t, the *atemporal content* of b is the proposition that p is true at t.

The atemporal content of my belief, formed in 2014, that Obama is the president of the United States is the (eternal) proposition that the proposition *that Obama is the president of the United States* is true in 2014. The atemporal content of my grandson's belief, formed in 2114, that Obama was not the president of the United States a century earlier is the (eternal) proposition that the proposition *that a century earlier Obama was not the president of the United States* is true in 2114. These atemporal contents are mutually incompatible and this seems to capture nicely the sense in which our two beliefs are at odds with one another: they put incompatible demands on the time-series.[37]

37 If it is supposed that the only propositions about Obama's presidency that my grandson and I believe are temporal, the atemporal contents of our beliefs should *not* be regarded as things that my grandson and I believe (or, for that matter, disbelieve). They should rather be taken as theoretical posits, useful to make sense of the idea of diachronic disagreement.

The suggestion is that similar moves can be made to account for the possibility of intersubjective disagreement. Suppose I BELIEVE *that Giovanni is special*, whereas, from your point of view, you BELIEVE *that it is not otherpersonally the case that Giovanni is special*. Intuitively, there is some kind of disagreement between you and me and this is so even if there is no point of view from which your belief and my belief are both BELIEFS. More importantly, this is so even if the propositions that your belief and my belief take as their object do not contradict one another. To recover the sense in which the two beliefs are in mutual disagreement, we need to look at their objective contents, where the objective content of any ordinary belief is defined as follows:

For any belief b in a proposition p, the objective content of b is the proposition that p is true from the point of view s from which b is a BELIEF.

A good criterion for intersubjective disagreement is:

(*Intersubjective Disagreement*) For any subjects x and y, x and y disagree if the objective content of x's beliefs is incompatible with the objective content of y's beliefs.

Besides being supported by the analogy with the temporal case, *Intersubjective Disagreement* is a good criterion for two reasons. First, it has the nice feature of entailing both *DISAGREEMENT* and *Disagreement** – an early indication that, by endorsing *Intersubjective Disagreement*, the friend of Subjectivism is on the right track to a coherent and unified account of disagreement.[38] Second, if *Intersubjective Disagreement* is a viable principle, a

38 *Intersubjective Disagreement* says that incompatibility of objective content is sufficient for disagreement. But whenever, from some point of view, two BELIEFS take incompatible propositions as their object there will be incompatibility of objective content. So *Intersubjective Disagreement* entails *Disagreement**. This, in turn,

grain of truth can be extracted from the simplistic view according to which *Disagreement* applies across the board to all beliefs. For take any purely objective proposition q (i.e. any proposition that, necessarily, is true from all points of view if it is true at all). And take any belief in q. Given that the proposition is objective, the objective content of that belief will be incompatible with the objective content of any belief in not-q, *no matter whether either or both beliefs are* BELIEFS. In other words: when q is a purely objective propositions, we can always be sure that a belief in q and a belief in not-q will amount to disagreement. And this is just to say that, if we restrict ourselves to purely objective propositions, *Disagreement* works just fine even if it is stated in terms of the ordinary notion of belief. The mistake of taking *Disagreement* to apply 'across the board" arose from ignoring the fact that, alongside purely objective propositions, there are subjective ones, and they, too, can give rise to disagreement.

If the issue is considered in the light of *Intersubjective Disagreement*, it's pretty clear that there need not be any disagreement between you and I if you believe yourself to be special and I believe myself to be special: we can both embrace the Subjectivist View of the Mental without disagreeing with one another.

Communication

Even if it doesn't lead to the conclusion that disagreement is utterly widespread, Subjectivism may be accused of making communication across different points of view impossible. As I sincerely report how things are from my point of view, I say something which may be false from your point of view. Symmetrically, as you sincerely report how things are from your point of view, you say something which may be false from my point of

entails DISAGREEMENT.

view. Perhaps we do not have to regard each other as irrational. Maybe we are not even disagreeing with each other. But how do we manage to communicate if neither of us can take on board the truth of what is said by the other, even if we are both speaking sincerely and mutually recognize that we are doing so?

I suspect that the answer will already be clear at this point, but it won't do much harm to outline it very briefly. In standard cases of communication, I take the utterances that I come across to be made by sincere and well-informed utterers. Sincere: I take the utterances to reflect the beliefs of their authors. Well-informed: I take the beliefs that the utterances reflect to be correct. So suppose that I come across an utterance of yours. I take that utterance to reflect a belief of yours and take that belief to be correct. If the utterance reflects a belief of yours, it reflects an otherpersonal BELIEF (for your beliefs are only otherpersonally BELIEFS). And if that belief is to be correct, what the utterance says has to be otherpersonally the case (this is what *Correctness** implies). So if you utter the words "Giovanni's mental states are not special" what I learn is not that Giovanni's mental states are not special (how could I? It is not true that Giovanni's mental states are not special!). What I learn is that *otherpersonally* Giovanni's mental states are not special. More precisely (given *Relative Correctness*), I learn that from your point of view Giovanni's mental states are not special. I learn something true, and I learn it on the basis of an utterance whose content is false.

There is nothing unique to intersubjective communication here – a similar dynamic can be observed in the cases of communication across time. Suppose you uttered the words "It's thundering" a moment ago. I take that utterance to reflect a belief of yours and take that belief to be correct. If the utterance was made a moment ago, it reflects a belief that you held a moment ago. And if that belief was correct, what the utterance says must have been the case a moment ago. So what I learn on the basis of your utterance is not that it's thundering (that might not even be the case anymore!). What I learn is that it used to be thundering. I learn

something true, but I learn it on the basis of an utterance whose content might very well be false.

So it's true that, if the world is not fully objective, when I hear your report on how the world is and you hear my report on how the world is, neither of us can always take on board the truth of what is said by the other – not even if, besides being well-informed, we are both speaking sincerely and mutually recognize that we are doing so. But, as the case of diachronic communication shows, it should never have been assumed that communication requires simply that one take on board the truth of whatever is said by one's interlocutor.

4. Subjectivism and Physicalism

So far, I've argued that SVM should not be regarded as a form of solipsism, that it is compatible with (a certain kind of) egalitarianism about points of view and that its defenders have all the resources to account for intersubjective phenomena like charitable interpretation, disagreement and communication. But a defense of SVM would hardly be satisfactory if it didn't include a discussion of the relationship between Subjectivism and Physicalism. In the last fifty years, much philosophy of mind in the analytic tradition has been haunted by “the metaphysical problem of explaining how our mentality is related to our physical nature” (Kim 1998, 1). What does SVM have to say about this problem? Does it make it unsolvable or more acute? Can SVM be reconciled with Physicalism?

Let us begin by stating the terms of the problem. To a first approximation, Physicalism is the thesis that everything is, or somehow 'reduces to', the physical. Various attempts have been offered to provide a more precise characterization, mostly in modal terms. Here I will adopt the formulation proposed by Jackson (1994, 28):

(*Physicalism*) Any world which is a minimal physical duplicate of the actual world is a duplicate simpliciter of the actual world.

A *physical duplicate* of the actual world is a world that is physically indiscernible from ours: think of it as a world that is exactly like ours insofar as physical entities and physical properties go. A *minimal* physical duplicate of the actual world is a physical duplicate of the actual world that does not contain any extra stuff (extra entities or extra properties). A duplicate *simpliciter* of the actual world is a world that is completely (and not just physically) indiscernible from ours. The thesis that Jackson identifies with Physicalism says that no minimal physical duplicate of the actual world fails to be a duplicate of it simpliciter.

While it may not be sufficient,[39] everyone agrees that accepting Jackson's thesis (or something along its lines) is at least necessary to qualify as a physicalist. And a good argument can be devised that, if one is a subjectivist, one cannot accept Jackson's thesis:

1) All the physical properties are objective properties.
2) So, if any world which is a minimal physical duplicate of the actual world is a duplicate simpliciter of the actual world, any world which is a minimal *objective* duplicate of the actual world is a duplicate simpliciter of the actual world.
3) But if Subjectivism is true, some worlds are minimal objective duplicates of the actual world without being duplicates simpliciter of the actual world.
4) So if Subjectivism is true, Physicalism is false.
   (by (2), (3) and Jackson's definition of *Physicalism*)

39 See Hawthorne (2002) for discussion.

Let me comment briefly on each of the premises of the argument.

I take (1) to be extremely natural and plausible. Suppose we call a property 'objective' if the proposition that an object has that property is a purely objective proposition (i.e. a proposition that, necessarily, is true from every point of view if it is true at all). The reason why it seems plausible to think that all the physical properties are objective properties is that the physical properties are nothing else than the properties studied by the natural sciences (better: by the final version of one particular science, physics) and the goal that the natural sciences set themselves in studying certain properties rather than others is that of providing an objective description of the world – a description whose truth, not varying from one point of view to another, could be endorsed by all of us, and potentially even by intelligent creatures radically different from us.

If all the properties recognized by physics are objective properties, then looking at a world that is exactly like the actual world in all the objective respects is looking at a world that is, *inter alia*, exactly like the actual world in all the respects recognized by physics. But Physicalism tells us that any world that is exactly like the actual world in all the respects recognized by physics is a perfect duplicate of the actual world (provided that it contains no extra stuff). It follows, *a fortiori*, that if Physicalism is true any world that is exactly like the actual world not only in all the respects recognized by physics but, more in general, in any objective respect, will be a perfect duplicate of the actual world (provided that it contains no extra stuff). So (2) is true.

The problem is that, if Subjectivism is true, the world is subjectively the way it is, which means that there is more to the world than what is objectively the case. Thus Subjectivism seems to entail that some worlds that agree with the actual world on all the objective truths disagree with it on some of the subjective ones. Supposing that at least some of these worlds contain no extra stuff , we arrive at (3): if Subjectivism is true, some minimal

objective duplicates of the actual world are not duplicates of it simpliciter.

(Caveat. I said that Subjectivism *seems* to entail that some worlds that agree with the actual world on all the objective truths disagree with it on some subjective ones. In reality, that is only so given further premises, in particular some principle to the effect that:

For every proposition p, if p is otherpersonally true then p is possibly true.

This principle requires defence and may be subject to counterexamples, depending on one's views about the interaction of subjectivity and modality. A detailed discussion of these issues would take us too far afield. For present purposes, I am happy to concede the truth of (3). It certainly seems natural and intuitive to think that *if* reality is subjectively the way it is *then* it could have been objectively just the way it actually is, contain no more stuff than it actually contains but have a different subjective character than the one it actually has).

Putting (1), (2) and (3) together, we get that Subjectivism and *Physicalism* are mutually incompatible: the first is false if the second is true and the second is false if the first is true. The argument seems to me to be valid and its premises plausible, so I don't think that subjectivists should resist its conclusion. What they can do is point out that, even if it is incompatible with the unrestricted version of physicalism defined by Jackson, Subjectivism can be reconciled with a slightly weaker physicalist thesis, according to which:

(*Objective Physicalism*) Any world which is a minimal physical duplicate of the actual world is an *objective* duplicate of the actual world.

If physicalism, as standardly defined, affirms that everything is, or 'reduces to', the physical, *Objective Physicalism* says that everything *objective* is, or 'reduces to', the physical. The

weakening is significant, but not unprincipled or unjustified. For one might think that, rather than signalling the failure of the physicalist project, the 'irreducibility' of the subjective to the physical is nothing else than the reflection of one particular constraint that the physical sciences decided to subject their inquiry to ("describe things as objectively as you can"). Moreover, *Objective Physicalism* is still a substantive thesis, and one that most, if not all, of today's self-proclaimed dualists would reject. (To mention a couple of examples, Chalmers's (1996) property-dualism about the phenomenal and Swinburne's (1986) substance-dualism about the mind are both incompatible with *Objective Physicalism*, as they both posit facts that, despite being objective, are irreducible to the physical ones).

This having been said, the point remains that (if the argument above is in good standing) subjectivists cannot endorse *Physicalism* as standardly conceived: they have to deny that a 'complete' description of the actual world can be given in purely physical terms. This is one of those things that are likely to make life difficult for the friends of Subjectivism, for "we live in an overwhelmingly physicalist or materialist intellectual culture [and] the result is that, as things currently stand, the standards of argumentation required to persuade someone of the truth of physicalism are much lower than the standards required to persuade someone of its negation" (Stoljar 2009). No argument or intuition in favour of Subjectivism will count for much as long as denying physicalism is regarded as 'the end of the world'.

But is it really 'the end of the world'? What are the alleged disastrous consequences of saying that not everything is, or 'reduces to', the physical? The word that is most often evoked in reply to this question is 'epiphenomenalism'.[40] Denying that everything is, or 'reduces to', the physical is supposed to imply that there are features of the world that are causally inert and this, in turn, is supposed to be bad or, at least, undesirable. But what's the argument from the falsity of physicalism to epiphenomenalism and why is epiphenomenalism so bad? Let me

40 See Papineau (2001) for a discussion of the relation between physicalism and the threat of epiphenomenalism.

address the two questions in turn.

There are many reconstructions of the argument from the falsity of physicalism to epiphenomenalism. The version that is supposed to make trouble for the friend of Subjectivism takes the following form:

(a) *Physicality*: Every caused fact has a physical cause (i.e. it is caused by a fact that is either physical or 'reducible' to a physical fact);

(b) *No (Massive) Overdetermination*: No caused fact (or nearly no caused fact) has two or more independent sufficient causes;

(c) *No Physicalism*: Subjective facts are neither physical nor 'reducible' to the physical facts. Hence:

(d) *Epiphenomenalism*: There are no facts (or nearly no facts) caused by the subjective facts.

There are a number of delicate issues surrounding the nature of causation. Here, in order to facilitate the discussion without trivializing matters, I will be making two simplifying assumptions that are supposed to make life easier for my opponent and more difficult for me:

- I am going to assume that causation is a relation among facts, and that there is more to this relation than the existence of certain regularities or the truth of certain counterfactual claims (denying this would be too easy a way of securing the causal efficacy of subjective facts: nothing prevents subjective facts from figuring in lawful regularities and worlds devoid of subjective facts are sufficiently remote in modal space to vindicate the truths of counterfactuals like "Had I not BELIEVED myself to be in danger, I wouldn't have left the place").[41]

41 For a discussion of how regularity- and counterfactual-based accounts of causation can be used to avoid the

- I will assume that we can make good sense of the notion of two 'independent' sufficient causes – in particular, I will be working on the assumption that cases of so-called 'causal inheritance' can be distinguished from cases of genuine overdetermination and be regarded as unproblematic cases of causation. To illustrate: the fact F *that the table is red* and the fact F' *that the table has surface spectral reflectance property P13* can both be causes of some further fact F" without this being a problematic case of overdetermination if the property of *being red* turns out to be a long disjunction of surface spectral reflectance properties or else a functional property of which P13 is the realizer (this much should be granted if we don't want the argument above to generalize to all the macroscopic facts that feature in the so-called special sciences – from biology and physiology to psychology and economics).

Now, suppose that, based on the argument above, we arrived at the conclusion that the subjective is entirely epiphenomenal: subjective facts are causally inert. Why would this be bad? For several different reasons. For one thing, you might agree with Samuel Alexander that something that "has nothing to do, no purpose to serve"- that is, something with no causal power- "might as well and undoubtedly would in time be abolished" (1920, 8).[42] Second, you might think that, being causally inert, subjective properties would be beyond the reach of semantic reference, because reference requires some kind of causal contact with the referent. Third, you might adopt the popular view that "[one] crucial factor in justifying a belief about an entity is an appropriate causal connection between the belief and the entity it is about" (Chalmers 1996, 193) and conclude, on such a basis, that if subjective facts were epiphenomenal nobody could justifiably believe in their existence.

Again, there are a number of moves open to the subjectivist that I will be ignoring

threat of epiphenomenalism, see Chalmers (1996, 151-152).

42 Cited by Kim (1997, 119).

here. Even if the subjective properties turned out to be causally inefficacious, it's not clear that we wouldn't be able to refer to them, for it's not clear that reference is *always* mediated by a causal contact with the referent (abstract objects, if there are any, represent one notable exception). It is also less than clear that every particular fact I believe in must be *causally* connected with my belief for the belief in question to be justified (epistemic justification could also be grounded by Russellian acquaintance, a relation that need not be construed as causal in nature). As to Alexander's thesis that what has nothing to do might as well be abolished, I see hardly any reason to accept it: not every difference needs to be a difference that makes a difference.

Fortunately, however, these moves are unnecessary. My preferred response to the charge of epiphenomenalism is to nip the argument in the bud by rejecting its chief premise, *Physicality*. All I think one should accept as a subjectivist (provided, of course, that one has prior sympathies for something along the lines of *Objective Physicalism*) is:

(*Objective Physicality*) Every caused *objective* fact is caused by a fact that is either physical or 'reducible' to a physical fact.

The difference between *Physicality* and *Objective Physicality* is crucial: it allows the subjectivist to hold that subjective facts have causal work to do, namely causing other subjective facts.

Let me give you an easy example. Take the subjective fact *that Giovanni is in* PAIN. It is perfectly compatible with *Objective Physicality* to hold that this fact is *not* causally inert, but, rather, is causally connected with the subjective fact *that Giovanni* BELIEVES *himself to be in* PAIN. (More importantly, it is compatible with *Objective Physicality* to hold that the two facts are causally connected in such a way as to account, at least in part, for the justification

of Giovanni's BELIEF or for the fact that Giovanni's BELIEF amounts to KNOWLEDGE). In brief, by accepting *Objective Physicality* instead of *Physicality*, subjectivists are in a position to say that subjective differences *do* make a difference: they make, first and foremost, a subjective difference.

Notice that saying that one subjective fact causes another is saying that it *subjectively* causes it. For here we are assuming causation to be a relation among facts and two facts can be causally connected only from a point of view from which they both exist. From your point of view, the fact *that Giovanni is in PAIN* does not cause the fact *that Giovanni BELIEVES himself to be in PAIN*, not least because, from your point of view, Giovanni is *not* in PAIN and does *not* BELIEVE to be in PAIN, so there aren't any such things as the fact *that Giovanni is in PAIN* or the fact *that Giovanni BELIEVES himself to be in PAIN*. This means that if subjective facts can be causes, at least some causal facts (i.e. facts as to what causes what) must be subjective rather than objective. Some might object to this, for causation is a physical relation and physical relations are objective relations (recall premise (1) of the argument for the incompatibility of Subjectivism and *Physicalism*). In line with *Objective Physicality*, my reply is that *objective* causation (i.e. causation objectively holding between objective facts) is a physical relation, but *subjective* causation (i.e. causation subjectively holding between subjective facts) need not be. This doesn't necessarily mean that subjective causation is causation of a radically different kind. It might simply mean that, occurring among subjective facts and holding subjectively rather than objectively, subjective causation does not fall in the domain of study of physics or any other natural science: it is not the kind of thing that the natural sciences would concern themselves with, but it can do many of the things that physical causation also does, in particular grounding epistemic justification and making semantic reference possible.

A more serious objection to allowing subjective facts to cause other subjective facts

arises within the framework of SVM. The worry is that, whenever a subjective fact causes another subjective fact, it will thereby cause an objective fact. But, given *Objective Physicality*, the objective fact in question will have a physical cause, too. Hence, we will have a case of overdetermination.

To illustrate, consider the fact *that Giovanni is in* PAIN and suppose that it causes the fact *that Giovanni* BELIEVES *himself to be in* PAIN. By causing the fact *that Giovanni* BELIEVES *himself to be in* PAIN, Giovanni's PAIN also causes the objective fact *that Giovanni believes himself to be in* PAIN. This follows from two simple assumptions. The first is the disjuctive definition of belief in terms of BELIEF:

To believe is to *either* BELIEVE *or* otherpersonally BELIEVE.

And the second is a principle of causal closure to the effect that:

(*Closure*) If A causes B and B realizes C then A causes C.

A fact *realizes* another when the latter holds in virtue of the former. Given the real definition of belief in terms of BELIEF (which is part of SVM), it seems natural to say that the subjective fact *that Giovanni* BELIEVES *himself to be in* PAIN realizes the objective fact *that Giovanni believes himself to be in* PAIN: the latter holds in virtue of the former. But then, by *Closure*, Giovanni's PAIN causes the latter by causing the former. This, in turn, might be thought to be in tension with *Objective Physicality* and *No (Massive) Overdetermination*. For if *Objective Physicality* is true, then the objective fact *that Giovanni believes himself to be in* PAIN is caused by a fact that is either physical or 'reducible' to a physical fact. And if *No (Massive) Overdetermination* is true, that must be the only cause of the fact *that Giovanni believes*

himself to be in PAIN.

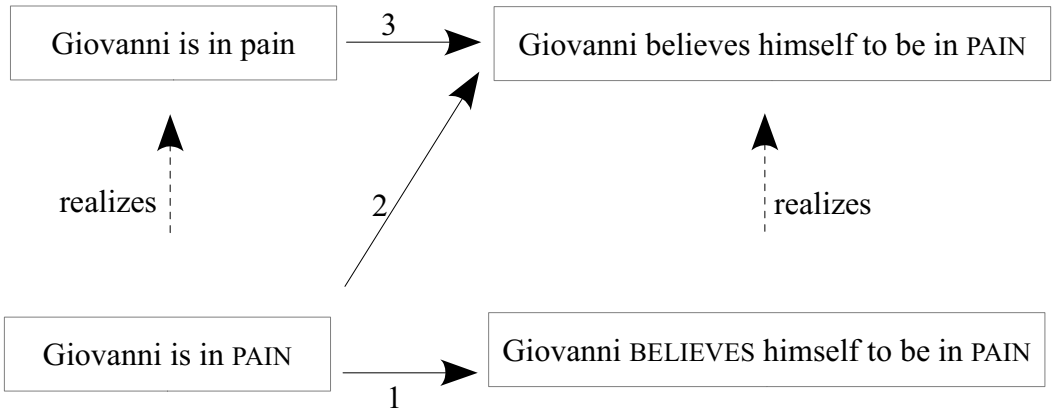
A subjectivist could reply by conceding that subjective causation generates *subjective* cases of overdetermination (Giovanni's belief is not *objectively* overdetermined, given that, from other points of view, it is not caused by the fact *that Giovanni is in* PAIN). But a much better reply is to deny that *No (Massive) Overdetermination* is violated. This can be done if the relationship between the subjective and the objective cause of the fact *that Giovanni believes himself to be in* PAIN is described as one of causal inheritance. Let me explain.

Suppose that the objective cause of Giovanni's belief that he is in PAIN has been identified: it is the fact *that Giovanni is in pain*. Suppose, further, that *Objective Physicalism* is true and the objective fact *that Giovanni is in pain* is, or 'is reducible to', a physical fact – so that *Objective Physicality* is satisfied. SVM has a familiar story to tell about the relationship between the objective fact *that Giovanni is in pain* and the subjective fact that *that Giovanni is in* PAIN. The story involves, of course, the disjunctive definition of pain in terms of PAIN:

To be in pain is to *either* be in PAIN *or* be an otherpersonally in PAIN.

Given this definition, it is natural to say that the fact *that Giovanni is in pain* is 'realized' by the fact *that Giovanni is in* PAIN, pretty much in the same sense in which the fact *that the table is red* is (on some accounts of what redness is) 'realized' by the fact *that the table has surface spectral reflectance property P13*. But, as I made clear at the outset, when two facts are so related, the observation that they both cause another fact can't be taken to be evidence of overdetermination (otherwise overdetermination would be found everywhere in the causal explanations of the special sciences). The right thing to say is, roughly, that the realized fact 'inherits' its causal efficacy from the realizer. This is what subjectivists should say about the

fact *that Giovanni is in pain* : it inherits its causal efficacy from the subjective fact *that Giovanni is in* PAIN. So the observation that they both cause Giovanni's belief to be in PAIN is not in tension with the truth of *No (Massive) Overdetermination*. The situation can be represented in the following way:



The dotted arrows indicate realization, whereas the numbered arrows indicate causation. There are three instances of causation. 1 is the most fundamental one and is a case of subjective causation (i.e. causation subjectively holding between subjective facts). 2 is dictated by *Closure* (Giovanni's PAIN causes Giovanni's BELIEF and whatever is realized by Giovanni's BELIEF). As to 3, it is an unproblematic case of causal inheritance (Giovanni's pain inherits its causal power from Giovanni's PAIN, by which it is realized).

The spirit of this strategy can be captured using a spatial metaphor. Most accounts that allow non-physical facts to do causal work get into trouble because they take the non-physical facts in question to be facts 'over and above' the physical facts (i.e. facts that neither realize nor are realized by the physical facts). This leads either to massive overdetermination or to violations of *Physicality*. But on the account proposed here, the relevant non-physical facts are not facts 'over and above' the physical facts: they are, rather, facts 'beneath and below' the physical facts. They *ground* (some) physical facts and it is *in virtue* of them that (some)

physical facts do the causal work that they do.

Let me sum up. It is true that Subjectivism and *Physicalism*, as standardly defined, are mutually incompatible. But Subjectivism is compatible with *Objective Physicalism*. Moreover, it doesn't follow from the fact that the subjective is neither physical nor 'reducible to' the physical that the subjective is causally inert. Subjective facts subjectively cause other subjective facts. They subjectively cause some objective facts too, but this doesn't give rise to any suspicious form of massive overdetermination. So subjectivists can avoid epiphenomenalism about the subjective and its (alleged) disastrous consequences for their view.

PART II

Most of us these days have learned to live with a certain picture of reality and of the place of the mental within it. In rough outline, the picture looks like this:

> Reality contains many things: mountains and rivers, plants and animals, stars and planets. Among the many things it contains, there are subjects who enjoy mental states: beliefs, hopes, desires, feelings and the like. Although these subjects resemble or differ from one another in many respects – including the mental states they enjoy – there is a clear sense in which they are all 'on a par': no subject is such that his or her mental states are, in any important respect, special or different from all the rest.

This picture – which may be called the Mainstream View of the Mental (MVM) – is *simpler* than SVM, because it doesn't require us to posit a plurality of points of view. And it's certainly more *egalitarian* than SVM, because it doesn't require us to draw any distinction between subjects with mental states and subjects with MENTAL states. Those who are altogether skeptical about the justificatory role of intuitions will be wondering what *reasons* I have for preferring SVM to MVM: intuitions aside, what's the evidence in favour of SVM and against MVM?

The second part of this thesis tries to provide an answer to this question, by showing that certain philosophical puzzles that arise within the framework of MVM can easily be solved (or dissolved) if one adopts SVM. These puzzles have to do with the asymmetry between self-knowledge and knowledge of other minds (Chapter 3), the immunity to error through misidentification of the judgments we form about our own mental states (Chapter 4)

and the essentially 'firsthand' character of experiential (or phenomenal) knowledge (Chapter 5).

As anticipated in the Preface, the discussion will be conducted in the spirit of integrating metaphysical and epistemological issues. The idea is that, insofar as it exhibits the peculiar features it exhibits, the epistemology of the mental confronts us with what, following Peacocke (1999), might be called an 'Integration Challenge':

> The problem is one of reconciliation. We have to reconcile a plausible account of what is involved in the truth of statements of a given kind with a credible account of how we can know those statements, when we do know them. [...] I call the general task of providing, for a given area, a simultaneously acceptable metaphysics and epistemology, and showing them to be so, the *Integration Challenge* for that area. (Peacocke 1999, 1)

In the present context, the challenge is that of explaining why our knowledge of mental facts has the particular character it has given certain hypotheses about the nature of mental facts. What I will argue is that the challenge is easier to meet for someone who accepts SVM (and, with it, the hypothesis that at least some mental facts are subjective rather than objective) than it is for someone who accepts MVM (and, with it, the idea that all facts – including mental facts – are objective).

CHAPTER 3:

Self-knowledge and knowledge of other minds

The idea that "a radical difference holds between our access to our own experience and our access to the experience of all other human beings" (Hyslop 2014) is probably as old as thinking about the mind. Some describe it as a deep-seated intuition (Bilgrami 2012, 263), others as a preconception shared by "a long and widespread philosophical tradition" (Schwitzgebel 2012a), yet others as something which forms "part of ordinary unphilosophical thought" (Wright 1998, 23). Whatever its genealogy, pretty much everyone agrees that there's something true to this idea. That my knowledge of my own mind is, in some sense, *unlike* my knowledge of other minds may not be a 'Moorean fact',[43] but it is certainly a very common view – one that we would hope to be vindicated by our best theories or, at least, not to be openly contradicted by them.

I will call this idea:

(*Asymmetry*) There is a radical difference between one's knowledge of one's own mental life and one's knowledge of the mental lives of others.

I will comment shortly on the sense in which the difference between self-knowledge and knowledge of other minds is (or is supposed to be) a 'radical' difference. Even before we come to that, however, one point should be apparent. While ordinary unphilosophical thought never

43 "One of those things that we know better than we know the premises of any philosophical argument to the contrary" to cite Lewis (1996, 549).

bothers itself with the issue, it is not philosophically obvious how *Asymmetry* can be made to square with MVM (i.e. the Mainstream View of the Mental) according to which:

(*Equality*) There is no radical metaphysical difference between one's own mental life and the mental lives of others.

MVM says that, metaphysically speaking, we subjects are all equal: no subject is such that his or her mental states are, in any important sense, special or different from the rest. But *Asymmetry* says that, epistemologically speaking, the mental realm is characterized by radical inequality: the kind of knowledge one has of one's own mental life is radically different from the knowledge one has of the mental lives of others. There seems to be some kind of mismatch here. Given *Equality*, one wouldn't expect *Asymmetry*. For if all subjects are on an equal footing, why should knowledge of what goes on with one particular subject (namely, oneself) be so deeply *unlike* knowledge of what goes on with all the other subjects? And given *Asymmetry*, one wouldn't expect *Equality*. For if knowledge concerning one particular subject (namely, oneself) is so different from knowledge concerning any other subject, isn't it tempting to think that there must be something special about that subject?

Now, since we've all learned to live with MVM, we've also learned a way of answering these questions that doesn't call *Equality* into question. The solution is familiar: it's true that self-knowledge differs radically from knowledge of other minds, but that's simply because we know about ourselves on the basis of introspection and we know about others by observing their behaviour. Metaphysically speaking we are all equal, but two radically different *ways of knowing* underpin self-knowledge and knowledge of other minds.

In this chapter, I will take a closer look at this familiar way of reconciling *Asymmetry* and *Equality*. I will try to trace back its origins, work it out in some detail and highlight its

costs and limitations (§ 2). Having done that, I will move to explain how one can make sense of *Asymmetry* if one abandons *Equality* – and with it MVM – to embrace the Subjectivist View instead (§ 3). One potential objection to the subjectivist account will then provide a natural starting point for discussing the relationship between *Asymmetry* and the so-called 'privacy' of the mental (§ 4).

1. Asymmetry

Before we start, some preliminary remarks are in order concerning *Asymmetry.*

First of all, *Asymmetry* speaks of a 'radical' difference between self-knowledge and knowledge of other minds. The adjective 'radical' is deliberately vague, for what the difference in question amounts to is part of what will be at stake in the rest of this chapter. However, at least three points should be relatively uncontroversial.

The first is that the difference *Asymmetry* is concerned with is not a matter of *degree*. The difference between a hot beverage and a very hot one is a matter of degree: a very hot beverage has to a higher degree or in greater amount what a hot beverage has to a lower degree or in lesser amount. In this sense, very hot beverages are not 'radically' different from hot ones. Now, perhaps some differences between self-knowledge and knowledge of other minds are a matter of degree (perhaps self-knowledge is more important for survival than knowledge of other minds – this difference is, or might be seen as, a matter of degree). But the point of *Asymmetry* is that there is at least one difference between self-knowledge and knowledge of other minds which is *not* like that: it is 'radical' in the sense that it cannot be

measured or represented on a scale.[44] In a word, it is a *categorical* difference, rather than a gradual one.

The second point is that the difference *Asymmetry* is concerned with is a pretty *natural* or 'joint-carving' difference. The difference between events that occurred before I brushed my teeth this morning and events that occurred after I brushed my teeth this morning is not very natural or 'joint-carving': a good taxonomy of the events in the history of the world would not give that difference much prominence. In this sense, the events that occurred before I brushed my teeth this morning are not 'radically' different from the events that occurred after I brushed my teeth this morning. Now, sure enough some differences between self-knowledge and knowledge of other minds are not very natural or joint-carving. But the point of *Asymmetry* is that there is at least one difference between self-knowledge and knowledge of other minds which (besides being categorical rather than gradual) is also very natural or joint-carving: it is 'radical' in the sense that it makes for real and genuine dissimilarity.

Third and last point: the categorical, joint-carving difference that holds between self-knowledge and knowledge of other minds does *not* hold between different instances of knowledge of other minds. If *Asymmetry* is true, a gulf separates my knowledge that I am in pain from my knowledge that Fred is in pain. But no analogous gulf separates my knowledge that Fred is in pain from my knowledge that Ted is in pain. So the difference *Asymmetry* is concerned with is 'radical' also in the sense that it is *unparalleled*: it can be found nowhere but in the comparison between self-knowledge and knowledge of any other mind.[45]

---

44 Saying that the difference between self-knowledge and knowledge of other minds is not a matter of degree is not quite the same as saying that there are no intermediate cases between self-knowledge and knowledge of other minds (although this is also true). Some differences are a matter of degree but do not admit of intermediate cases. Arguably, the difference between having 300 hairs and having 301 hairs is like that.

45 Here and in what follows, "self-knowledge" refers to knowledge of one's *current* mental life and "knowledge of other minds" refers to knowledge of the *current* mental life of other subjects. Arguably, there is a radical difference also between my knowledge of my future (or past) mental life and my knowledge of the future (or

Asymmetry says that self-knowledge and knowledge of other minds are radically different (where this implies, at least, a categorical, joint-carving, unparalleled difference). There is a question concerning the modal status of this principle: is it contingent or necessary?

Tradition has it that the radical difference between self-knowledge and knowledge of other minds is not a sheer contingency. Alston (1971) notes that according to many philosophers it is a *criterion* for the mental or the psychological that the state of a person is mental or psychological if and only if that person's knowledge that she is in that state is radically different (in the relevant sense) from the knowledge of that fact that is available to anyone else. This view implies that *Asymmetry* is necessarily true.[46] Of course, tradition can be wrong. Maybe there are metaphysically possible worlds in which *Asymmetry* is false. What seems plausible is that, if there are worlds in which *Asymmetry* is false, they are rather remote worlds (at least if we take distances among worlds to reflect their similarity, as is usual). For example, a world in which self-knowledge and knowledge of other minds are pretty much 'on a par' is not as close to the actual world as one in which yesterday I wore blue trousers instead of black ones or in which life did not develop on Earth. The thought is that, even if it is only contingently true, *Asymmetry* couldn't have *easily* been false. More likely than not, *Asymmetry* is a feature of a fairly large range of worlds surrounding the actual one.

A final reflection concerns the distinction between *Asymmetry* and another thesis, that I shall call:

(*Superiority*) One's knowledge of one's own mental life is *superior* to one's knowledge of the

---

past) mental life of other subjects – I will remain neutral on this point.

46 McGinn (2004) discusses a nice contrast case: our epistemic access to near objects is different from (and by and large better than) our epistemic access to far objects, but this strikes us as a mere contingency: "we can easily imagine that our senses were sensitive in the opposite direction, so that we are better turned to the properties of far objects than near ones" (2004, 238). The standard view is that this is not so with the difference between self-knowledge and knowledge of other minds.

mental lives of others.

Superiority says that self-knowledge is not only different, but also superior to knowledge of other minds. Whether or not this claim entails *Asymmetry* depends on how we gloss the notion of one kind of knowledge being "superior" to another. If the superiority in question is merely a matter of degree and pertains to features that are not very natural or 'joint-carving', *Superiority* does not entail *Asymmetry*. On the other hand, if the idea is that self-knowledge is 'radically' superior to knowledge of other minds, *Superiority* entails *Asymmetry*, but is not entailed by it.

The notion that we have better access to our own mind than to the minds of other subjects is a recurrent motif in the literature on self-knowledge and goes back at least as far as Descartes, who put it by saying that I "know plainly that I can achieve an easier and more evident perception of my mind than of anything else" (Descartes 1988, 86). It is far from clear that Descartes had in mind a 'radical' kind of superiority, though. For example, he did not seem to think that we are infallible or omniscient about our own minds (his view was that "we frequently make mistakes even in our judgments concerning pain" (Descartes 1988, 183)).[47] What is true is that many philosophers who identify themselves with the "Cartesian" tradition have explicitly defended claims of 'radical' superiority. The varieties of 'radical' privileged access abound,[48] but one thesis that has often been discussed in this connection is that we have privileged access to our own mind, because mental states (or, at least, a suitably large class of mental states) are *luminous*. For a state to be luminous is for the condition that one is in that state to be luminous, meaning that one is always in a position to know that one is in that condition whenever one is in that condition.[49] To illustrate, suppose that pain is a luminous

47 See Newman (2010) for discussion.

48 Alston (1971) provides an exhaustive taxonomy. See also Gertler (2011).

49 More carefully, a condition C is *luminous* if and only if, for every case α, if C obtains in α, then one is in a

state. Then we are in a position to know that we are in pain whenever we are in pain: the fact is open to our view, unhidden, and with no obstacle to our knowing it. If mental states (or, at least, a suitably large class of mental states) are luminous, a 'radical' version of *Superiority* is vindicated. For certainly facts concerning the mental states of other subjects are *not* open to our view, unhidden and with no obstacle to our knowing them. And, arguably, this makes knowledge of other minds 'radically' inferior to knowledge of any luminous condition.

In recent years, the thesis that mental states are luminous – and with it, many other attempts to capture the sense in which self-knowledge is 'radically' superior to knowledge of other minds – has come under considerable attack. Williamson (1996; 2000) argued that the only luminous conditions are the trivial ones (i.e. those that obtain in every case whatsoever or in none at all). His argument exposes a tension between the thesis that there are non-trivial luminous conditions and the plausible assumption that knowledge requires reliably based confidence. Since, for any mental state m, being in m is clearly not a trivial condition, it follows that mental states are not luminous. Conee has offered what he regards as a better argument against the luminosity of any condition whatsoever: to be in a position to know that p, "no obstacle must block one's path to knowing that p" (Williamson 1996, 555), but one's path to knowing that p can always happen to be blocked by credible, even if misleading, evidence to the contrary (e.g. the testimony of an expert), so no condition is luminous (Conee 2005, 448-449). More recently, Greenough (2012) has tried to show that the difficulties Williamson raised for the thesis that mental states are luminous arise also for the thesis that mental states are *lustrous* – where a state is lustrous if and only if one cannot be in that state without being in a position to justifiably believe that one is in that state (Berker 2008).

Anti-luminosity and anti-superiority arguments have stirred lively reaction. The

position to know that C obtains in α (see Williamson 1996). Think of a case as a triple, <w, s, t>, consisting of a world, a subject and time.

discussion is vast and intricate and the jury is still out on many issues. Rather than entering that debate, in what follows I will focus my attention on *Asymmetry* and say little or nothing about *Superiority*. The rationale for this choice is twofold. First of all, it seems to me that, if something along the lines of *Superiority* is true, the explanation of why this is so would have to involve a careful study of *Asymmetry*. A proper understanding of the way (or ways) in which self-knowledge differs from knowledge of other minds might illuminate the way (or ways) in which self-knowledge is superior to knowledge of other minds (if it really is). *Asymmetry* is explanatorily more basic or fundamental than *Superiority*, or so it seems to me. Secondly, one respect in which our access to our own mind might be described as 'superior' to our access to other minds will be discussed in the next chapter, where I will look at why certain judgments we make about ourselves, unlike analogous judgments we make about others, exhibit what is sometimes called 'immunity to error through misidentification'.

2. *Asymmetry* and the Mainstream View of the Mental

The Mainstream View of the Mental holds that, metaphysically speaking, my mental life is not radically different from the mental lives of all human beings. Why, then, is my access to my own mental life so radically unlike my access to the mental lives of other human beings? The answer to this question is not univocal: defenders of the Mainstream View have defended a variety of different approaches to *Asymmetry*. The one I will focus on here – which is probably the most popular and historically most influential – rests on the idea that self-knowledge and knowledge of other minds result from different *ways of knowing*.[50]

50 Let me mention in passing three mainstream accounts of *Asymmetry* that I will not consider. According to the first, it is built into the linguistic rules governing the use of mental predicates (or it is a consequence of those

Before entering the details of the proposal, we need to say a little about what a *way of knowing* is. First of all, a few examples. My way of knowing that there is a table in front of me is by seeing it. My way of knowing that a person has entered the room is by hearing that person entering the room. More in general, perception is a way of knowing. So is testimony: for example, one can come to know what time it is by asking someone else and one can learn that Caesar crossed the Rubicon in 49 BC by reading history books. Another way of knowing is deduction: I can know that 8 is divisible by 2 by deducing it from the fact that 8 is even and even numbers are divisible by 2. What do all these examples have in common?

According to the account I will be assuming here, which is the one advocated by Cassam (2007a), the essential feature of a way of knowing is that it is suitable to feature in a satisfactory *explanation* of knowledge. More precisely, Φ-ing is a way of knowing that p only if "it is possible satisfactorily to explain how S knows that p by pointing out that S Φs" (Cassam 2007a, 340). One feature of this account is that it is fairly permissive. As Cassam notes, a way of knowing need not be a relation to a proposition (seeing the table in front of me is a way of knowing that there is a table in front of me, but "seeing" does not express a relation to a proposition in this context). And even in those cases where they are relations to propositions, ways of knowing need not be factive relations (reading that p is a way of knowing that p, but does not entail that p). A fortiori, it is not true that, if F-ing that p is a way

---

linguistic rules) that something along the lines of *Asymmetry* holds – Fricker (1998) calls this the "artefact of grammar" theory while Wright (1998) labels it the "default view". According to the second, *Asymmetry* holds because knowledge of other minds requires a substantive cognitive effort, whereas self-knowledge is merely a by-product of our customary knowledge of the external world (on one way of developing this view, "I get myself into the position to answer the question whether I believe that *p* by putting into operation whatever procedure I have for answering the question whether *p*" (Evans 1982, 225)). This is what Schwitzgebel (2012a) calls the "transparency" account of self-knowledge. According to the third account, self-knowledge and knowledge of other minds are involved with substantially different modes of presentation of mental properties: self-knowledge, unlike knowledge of other minds, requires the exercise of essentially 'experiential' modes of presentation. I will say more about experiential modes of presentation in Chapter 5.

of knowing that p, F-ing that p entails knowing that p.

(Cassam's proposal contrasts sharply with the account of ways of knowing proposed by Williamson (2000), on which ways of knowing are factive attitudes and bear to knowledge the relation that a determinate bears to its determinable. On Cassam's conception, ways of knowing are not so much specific modes of knowing things as specific ways of coming to know things.[51] My reasons for favouring this account over Williamson's are exquisitely dialectical. One kind of question that will be relevant in what follows is the question, concerning some specific x, whether x is a genuine way of knowing things. I do not want the answer to that question to be sensitive to whether x is a propositional attitude, is factive or entails knowledge. All I require of a genuine way of knowing things is that it be suitable to feature in a satisfactory explanation of how one can come to know things. While I am fairly liberal about what counts as a satisfactory *explanation* of how one can come to know things, my liberalism does not go as far as allowing that, trivially, whenever there are distinct pieces of knowledge there are also distinct ways of knowing in terms of which they can be explained. I take it that there are intuitive requirements that any genuine *explanans* of knowledge should satisfy – the examples above are supposed to provide the reader with a rough sense of what these requirements are).

Let us now go back to the Mainstream View of the Mental and its take on *Asymmetry*. The proposal I want to focus on has it that self-knowledge and knowledge of other minds are radically different because self-knowledge is gained through *introspection*, whereas knowledge of other minds is gained through an *inference from observed behaviour*. Introspection and inference from observed behaviour are supposed to be two different ways of knowing, which explain (respectively) how we come to know about our own mental life and

51 Cassam draws a subtle distinction between ways of knowing and ways of coming to know (Cassam 2007a, 350-351), but the difference won't matter too much for my purposes.

about the mental life of others. As we shall see, the proposal calls for some qualifications, but first of all let us lay out its basics.

2.1. Introspection

We can start with *introspection*. It is standard to characterize introspection as a way of knowing whereby one gains direct knowledge about one's mental properties (Schwitzgebel 2012a; Gertler 2011). Introspection allows one to know only about one's *mental* properties: I cannot know my address on the sole basis of introspection. And it only allows one to know about *one's own* mental properties: "any process that generates knowledge equally of one's own and others' minds is by that token not an introspective process" (Schwitzgebel 2012a). Furthermore, it is supposed to provide one with *direct* knowledge of one's mental properties. This is a bit more difficult to explain. Traditionally, the directness in question has been construed as both metaphysical and epistemological. Introspection is a metaphysically direct way of knowing in the sense that the knowledge it yields is directly (or immediately) caused by the fact it is knowledge of: pain is the direct (or unmediated) cause of my introspective knowledge that I am in pain. It is an epistemologically direct way of knowing in the sense that it does not involve any conscious or unconscious inference from a set of premises to a conclusion, in much the same way as visual perception does not involve any conscious or unconscious inference from a set of premises to a conclusion.[52] Introspection is not total: I am not currently introspecting all of my mental states. Nor is it constant: one exercises it at some

52 There is a stronger sense of epistemological directness on which for a way of knowing to qualify as epistemologically direct is for it to yield knowledge without the causal mediation of any other piece of knowledge. Although some have defended the claim that introspection is direct in this stronger sense, the resulting position is rather extreme and I will not discuss it here.

times and not others. In this sense, it should not be confused with what is sometimes called 'direct awareness' or 'acquaintance' – a relation that is supposed to hold *at all times* between a subject and *all* his or her experiences.[53]

The history of the notion of introspection is long, but not as long as one would expect. In ancient philosophy, for instance, self-knowledge is not associated with a distinctive way of knowing (Woolf observes that "Plato lacks the Cartesian notion of introspection, with its accompanying doctrine of privileged first person access" (2008, 94)). Gilbert Ryle puts forward the hypothesis that the notions of consciousness and introspection originated only in the modern age, in part as "a transformed application of the Protestant notion of conscience" and in part as "a piece of para-optics" inspired by the study of light in Galileian science (Ryle 1949, 159). The hypothesis is intriguing, but unconvincing. We know that "taking their cue from Augustine's account of self-knowledge, medieval philosophers [held] that knowledge regarding our own mental states is epistemically distinctive in a number of ways" (Brower-Toland 2012, 1). They also had a lively debate as to whether or not this epistemic distinctiveness should be explained in terms of some kind of "reflexive intellective intuitive cognition" – a way of knowing ('cognition') whereby the minds (or 'intellect') acquires direct (or 'intuitive') knowledge of its own states ('reflexive').[54]

Ryle's suggestion is not completely off track, however. It's true that it is only in the modern age that the idea of explaining the unique status of self-knowledge in terms of a distinctive way of knowing became a dominant research paradigm in epistemology. Malebranche distinguished four ways of perceiving or knowing things: by themselves,

53 The notion of acquaintance was first introduced by Russell (1910). It plays a crucial role in the work of Chalmers (2003a). For a critical discussion of this notion, with which I am very sympathetic, see Hellie (2013).

54 Brower-Toland (2012) focuses on self-knowledge in Ockham and Chatton. For a discussion of Aquinas's view that "the intellect 'perceives' its activities, and that it is from this perception that all self-knowledge arises" see Kenny (1993, 122ss).

through their ideas, through consciousness or inner sensation and through conjecture – the third being, of course, the way we gain knowledge about our own soul and "what is taking place in us" (Malebranche 1997, 237). The same idea can be found in Locke, who famously identified a "source of ideas every man has wholly in himself; and though it be not sense, as having nothing to do with external objects, yet it is very like it, and might properly enough be call'd internal sense. […] I Call this *reflection*, the ideas it affords being such only as the mind gets by reflecting on its own operations within itself" (Locke 1975, 105). The label of 'inner sense' was then borrowed by Kant, who used it to designate a property of our mind "by means of which the mind intuits itself, or its inner state"(Kant 1998, 157) – "intuits" being Kant's term for a type of cognition that relates directly or immediately to its objects.

In the contemporary debate, the idea of introspection has been employed in a variety of different ways. Some have proposed to identify introspection with a quasi-perceptual mechanism, or cluster of mechanisms, allowing the brain to detect and keep track of its own states.[55] Others have emphasized the differences between introspection and perception, insisting that, through introspection, the mind not only knows itself, but shapes itself in distinctive ways.[56] Yet others have hypothesized the existence of 'constitutive links' between introspection and the mental states we know through introspection.[57] We don't need to go into these details here. The common theme I will be concerned with is the existence of a way of knowing whereby one gains knowledge of *mental* facts (rather than facts of other sorts), the knowledge one gains concerns exclusively *one's own* mental properties and has the distinctive feature of being metaphysically and epistemologically *direct*.

---

55 This is the view originally put forward by Armstrong (1968; 1981).

56 See, for instance, McGeer (1996).

57 Different versions of this idea can be found in the work of Tyler Burge (1988) and Sydney Shoemaker (1994). In recent years, the idea has been applied mainly to phenomenal introspection, i.e. introspective knowledge of one's experiences – see, for instance, Gertler (2001), Papineau (2002), Chalmers (2003), and Horgan and Kriegel (2007).

2.2. Inference from observed behaviour

On the mainstream account we are considering, introspection contrasts with the way of knowing that explains knowledge of other minds: we know ourselves through introspection and we know other minds through an inference based on how their possessors behave. Inference is a paradigmatically *indirect* way of knowing. It is metaphysically indirect because the knowledge it generates need not be immediately caused by its objects (I can infer that some remote planet exist without there being any direct causal link between my knowledge and the existence of that planet). And it is epistemologically indirect because it involves some kind of conscious or unconscious transition from a premise or a set of premises to a conclusion. In the case at hand, the inference is supposed to be from premises concerning observed behaviour: I see (or otherwise know) that someone is behaving in a certain way, I know that person to be relevantly similar to me, so I infer that his or her behaviour reflects a mental state similar to the one I am in when I behave in that way (the inference is justified by the principle that similar effects have similar causes, or by some other analogical principle of that sort).

Once again, the idea has a long history, although a good case can be made that it is foreign to ancient philosophy. In the ancient world, knowledge of other minds was not thought of as (epistemologically or metaphysically) indirect. For example, it has been argued that, according to Plato, other minds lend themselves to direct inspection and "are to be regarded as no more opaque to examination than other bodies" (Woolf 2008, 95). This might reflect the fact that the mind "in antiquity has a shared and universal side [instead of being] cut off from the world, and hence private and potentially inaccessible to others" (Remes 2008, 157). It is difficult to pin down when exactly things started to change, but it is clear that early Christian philosophy presents us with a rather different picture of our epistemological relation

to others. For example, Augustine may not have fully endorsed the idea that we know other minds on the basis of an inference from premises about observed behaviour, but it is difficult to deny that the spirit of the idea is already present in the pages of *De Trinitate*.[58] Interestingly, the rise of the inferential account of knowledge of other minds coincides with the rise of the introspective account of self-knowledge. As we've already seen, Malebranche distinguished four ways of knowing things. One of them ('through consciousness or inner sensation') corresponds to what I here call 'introspection'. Another is knowledge through conjecture, which is the only way we have of knowing the "souls of other men": "we do not know them either in themselves or through their ideas, and as they are different from ourselves, we cannot know them through consciousness. We conjecture that the souls of other men are of the same sort as our own. We suppose them to feel what we feel in ourselves" (Malebranche 1997, 239). Centuries later, the same idea is defended at great length by Mill, who presents it as follows:

> By what evidence do I know, or by what considerations am I led to believe [...] that the walking and speaking figures which I see and hear, have sensations and thoughts, or in other words, possess Minds? […] I conclude that other human beings have feelings like me, because, first, they have bodies like me, which I know, in my own case, to be the antecedent condition of feelings; and because, secondly, they exhibit the acts, and other outward signs, which in my own case I know by experience to be caused by feelings (Mill 1872, 243)

58 "For even when a living body is moved, there is no way opened to our eyes to see the mind, a thing which cannot be seen by the eyes; but we perceive something to be contained in that bulk, such as is contained in ourselves, so as to move in a like manner our own bulk, which is the life and the soul" (Augustine 1974, 120). But see Avramides (2001, 47ss) for an argument that this is not an instance of a 'reasoning from analogy' for the existence of other minds. Cary (2000) argues that it is Augustine who invented the concept of the self as a private inner space.

In the contemporary debate, the inferential account is still very influential. Hill (1991) explicitly endorses and defends Mill's view, on which we know about other minds on the basis of a conscious inference. This inference takes as its premises the behaviour that other subjects display and the assumption that other subjects are relevantly similar to ourselves (Hill 1991, 212). Other approaches allow for the possibility of treating the relevant inference as unconscious. For instance, according to some versions of the so-called 'theory-theory' of mentalizing, it is true that we ascribe mental states to others on the basis of a theoretical inference, but the major premise of this inference – a naïve psychological theory that we develop and adjust over time in early childhood – is known by us only unconsciously (or 'implicitly' or 'tacitly').[59] According to other models, the theory of mind on the basis of which we ascribe mental states to others is known by us not only unconsciously (or 'implicitly' or 'tacitly') but also innately: there is a 'theory of mind mechanism' that we are biologically equipped with.[60] The variants are many and the differences subtle. For present purposes, I will focus on what can be seen as the theme common to all these proposals, namely the existence of a way of knowing whereby one gains knowledge of *other minds* and the knowledge one gains is metaphysically and epistemologically *indirect*.

If we know about our own mental life on the basis of introspection and we know about the mental lives of other subjects on the basis of an inference from premises about observed behaviour (hereafter, inference from observed behaviour), the difference between self-knowledge and knowledge of other minds is at least as 'radical' as the difference between introspection and inference from observed behaviour. But the latter difference is no doubt pretty 'radical': it is not a matter of degree, it makes for a fairly deep (or 'joint-carving')

59 This is the so called 'theory-theory' of mentalizing. See Gopnik and Wellman (1992; 1994) and Gopnik, Meltzoff, and Kuhl (1999).

60 Leslie (1987).

epistemological distinction and it can be found nowhere but in the comparison between self-knowledge and knowledge of other minds. So by saying that self-knowledge and knowledge of other minds are acquired through these different ways of knowing, an advocate of the Mainstream View of the Mental seems well-positioned to vindicate *Asymmetry.*

Notice that the proposal need not be that we acquire self-knowledge *only* through introspection and that we acquire knowledge of other minds *only* on the basis of an inference from observed behaviour. It can equally well be that the way of knowing *typical* of self-knowledge is introspection, while the way of knowing *typical* of knowledge of other minds is inference from observed behaviour. The idea is that a way of knowing can be typical of knowledge of a certain kind even if knowledge of that kind is not *always* to be explained in terms of the way of knowing in question (just as barking can be typical of dogs even if not *all* dogs bark). In fact, one could be even less committal and vindicate *Asymmetry* by endorsing either (but not both) of the following two claims:

(*Introspection*) There is a radical difference between self-knowledge and knowledge of other minds because the way of knowing typical of self-knowledge is introspection and introspection is 'radically' different from whatever way of knowing is typical of knowledge of other minds.

(*Inference*) There is a radical difference between self-knowledge and knowledge of other minds because the way of knowing typical of knowledge of other minds is inference from observed behaviour and inference from observed behaviour is 'radically' different from whatever way of knowing is typical of self-knowledge.

It is these two (relatively weak) claims that I will take issue with in the rest of this section.

2.3. Critique

I will start from an objection that has to do with the modal status of *Asymmetry*. We've seen that, traditionally, *Asymmetry* has been taken to be necessarily true (§ 1). But an argument can be made that *Introspection* and *Inference* yield only a contingent version of *Asymmetry*. The chief premise of the argument is:

> (1) For any object of knowledge x and any way of knowing y, there are no necessary connections between (knowledge of) x and y;

The argument, then, has two distinct prongs:

> (2) One's mental life is an object of knowledge;
> (3) Introspection is a way of knowing;
> (4) So there are no necessary connections between knowledge of one's mental life and introspection.

> (2') Other minds are an object of knowledge;
> (3') Inference from observed behaviour is a way of knowing;
> (4') So there are no necessary connections between knowledge of other minds and inference from observed behaviour.

It is clear how (4) and (4') make trouble for someone who takes *Asymmetry* to be a necessary truth and wants to use *Introspection* and *Inference* to explain why this is so. If the radical difference between self-knowledge and knowledge of other minds is to be explained in terms

of a difference between ways of knowing (as the two principles say) and the latter difference is only contingent (as (4) and (4') suggest), it can't be but a mere contingency that self-knowledge and knowledge of other minds are radically different. And so *Asymmetry* can't be a necessary truth, after all.[61]

An argument along these lines is offered by McGinn (2004). McGinn defends (1) by saying that it "reflects the familiar realist separation between ontology and epistemology: the object of knowledge is not to be conflated with the knowledge itself" (2004, 237). This general principle is then applied to the particular case of self-knowledge and knowledge of other minds. McGinn is happy to concede that there is a way of knowing typical of self-knowledge (i.e. introspection) and a way of knowing typical of knowledge of other minds (i.e. inference from observed behaviour). His point is just that these ways of knowing are not *essential* to (respectively) self-knowledge and knowledge of other minds. There must therefore be worlds in which self-knowledge and knowledge of other minds are *not* underpinned by radically different ways of knowing.

McGinn offers two examples of how this might go. In the first example, we "rig up the world in such a way that facts about physical objects [and other minds] are immediately fed into a person's brain and trigger beliefs in their existence, without any perceptual mediation". The idea is that "we could describe this as a case in which either the subjects are "introspecting" external facts or have a mode of access to them that mirrors introspection in central respects" (McGinn 2004, 241). In the second example, we rig up the world in such a way that a person's knowledge of his or her own mental states is obtained only on the basis of

61 In principle one could object to this line of reasoning by pointing out that, even if there are possible worlds in which self-knowledge and knowledge of other minds are *not* gained through radically different ways of knowing, the worlds in question are ones in which self-knowledge and knowledge of other minds are radically different for other reasons (i.e. reasons *not* having to do with ways of knowing). But it seems plausible to think that, if *Asymmetry* holds true in every possible world, it holds true for the same reason in every possible world.

behavioural criteria. The idea is that we could describe this as a case in which the subject has a mode of access to his or her own mental life that mirrors knowledge of external facts (including other minds) in central respects. McGinn goes some way towards detailing the two scenarios, suggesting that both are metaphysically (if not also nomologically) possible and that both describe situations in which no 'radical' difference holds between the way of knowing typical of self-knowledge and the way of knowing typical of knowledge of other minds.

I do think that arguments of this kind put some pressure on accounts of *Asymmetry* based on *Introspection* and *Inference*, but I don't take it to provide decisive evidence against them.[62] I suspect that one can accept the realist separation between ontology and epistemology without accepting the idea that there are no necessary connections between domains of knowledge and ways of knowing.[63] Moreover, even if there are no necessary connections between domains of knowledge and ways of knowing and the conclusion of the argument is granted, this does not show that *Asymmetry* could have *easily* been false. It could still be the case that the worlds described by McGinn (i.e. worlds in which self-knowledge and knowledge of other minds are not 'radically' different) are extremely remote from the actual world. In that case, *Asymmetry* wouldn't be necessarily true, but it would still express a deep truth about the world we live in.

There are, however, more direct objections to *Introspection* and *Inference*. For *Introspection* to be true, introspection must be *the* way of knowing typical of self-knowledge. And for *Inference* to be true, inference from observed behaviour must be *the* way of knowing is typical of knowledge of other minds. Now, forget the question whether either or both of these theses are necessarily true (or true in a large range of worlds surrounding the actual

62 For a discussion, see Cassam (2004).

63 This is especially so if we are liberal about ways of knowing. See above (§ 2).

one). Is either of them even just contingently true? Is it true that, in the actual world, there are is a *single* way of knowing typically associated with self-knowledge and a *single* way of knowing typically associated with knowledge of other minds? Despite the long and venerable tradition supporting *Introspection* and *Inference*, it is far from clear that the answer to these questions is 'yes'.

Let us start by considering *Introspection*. There are several grounds for denying its truth. We can start from the extreme view that introspection – defined as a way of knowing whereby one gains (metaphysically and epistemologically) direct knowledge about one's mental properties – does not exist. Introspection, on this view, is a philosophical myth: no way of knowing satisfies the basic criteria that a way of knowing should satisfy in order to qualify as a kind of introspection. Some ways of knowing (e.g. perception) yield metaphysically and epistemologically direct knowledge, but the knowledge they yield is not about one's mental properties. Other ways of knowing (e.g. inference from observed behaviour) yield knowledge of mental properties, but they are neither metaphysically and epistemologically direct nor dedicated to the mental properties of just one subject. No way of knowing is *both* metaphysically and epistemologically direct *and* conducive to knowledge that concerns specifically one's own mental properties. So no way of knowing deserves the label of 'introspection'.

This extreme line is not without defenders. According to some brands of behaviourism, observation of our own behaviour is pretty much all we have to go by when deciding what mental states we are in – in this respect, self-knowledge is not only (metaphysically and epistemologically) indirect, but also perfectly 'on a par' with knowledge of other minds. A version of this thesis was advocated by Ryle, who famously dismissed "introspection" as "a term of art and one for which little use is found in the self-descriptions of untheoretical people" (Ryle 1949, 152). But we need not go as far as accepting

behaviourism to cast doubt on the existence of introspection. As Nichols and Stitch point out, "the basic idea of [the account of self-awareness that is most widely held among psychologists] is that one's access to one's own mind depends on the same cluster of cognitive mechanisms that plays a central role in attributing mental states to others" (Nichols and Stitch 2004, 297).[64] Suppose this is right and suppose there is a close connection between ways of knowing and the "cluster of cognitive mechanisms" on which they depend (as seems plausible). Then whatever way of knowing is associated with self-knowledge is also associated with knowledge of other minds. And this is just to say that there is no such thing as introspection.

I won't here elaborate on virtues and vices of this extreme line. I myself find it rather attractive, but, for present purposes, I am prepared to concede that it is misguided. What I want to do is offer my reader a more moderate set of reasons to doubt the truth of *Introspection*. The moderate line I have in mind goes as follows. First of all, there is not just one, but *many* ways of knowing whereby one gains metaphysically and epistemologically direct knowledge of one's mental properties.[65] In this sense, "introspection" is not so much a non-referring term as a term that does not designate a single epistemological natural kind.[66] Furthermore, while there are many ways of knowing corresponding to talk of "introspection", none of them is *typical* of self-knowledge. Rather, self-knowledge is gained in a variety of different ways, none of which is largely or significantly predominant. If the moderate line is correct, it is simply not true that there is a way of knowing typical of self-knowledge and 'radically' different from the way of knowing typical of knowledge of other minds.

64 For a recent defense of this thesis, see Carruthers (2011).

65 At least if we are not unreasonably demanding about what should count as metaphysically and epistemologically *direct* knowledge.

66 Compare: "jade" is not a non-referring term, but a term that does not designate a mineral natural kind – every instance of jade being either nephrite or jadeite.

Epistemologically speaking, self-knowledge is simply not a unitary category: *no* way of knowing (whether introspective or not) is typical of it.

An objector might say that the claim that *no* way of knowing is typical of self-knowledge does not sound 'moderate' at all. But notice that the claim only makes sense against the backdrop of the general account of ways of knowing proposed at the beginning of this section. It will be recalled that for something to be a way of knowing, it must be possible satisfactorily to *explain* knowledge in terms of it. Now, the moderate approach I've just sketched is compatible with the claim that there is a certain gerrymandered bunch of activities we engage in whenever we want to attain self-knowledge. And it is also compatible with the claim that it is only to attain self-knowledge that we engage in that particular gerrymandered bunch of activities – other bunches of activities being required for knowledge in other domains. All the moderate approach says is that there is no single *way of knowing* corresponding to the bunch of activities we engage in to attain self-knowledge. This is just to say that a good explanation of self-knowledge will appeal to different methods or processes in different cases. In a slogan: self-knowledge has no single explanans and, therefore, no dedicated way of knowing.

The moderate view of self-knowledge has been recently defended by Schwitzgebel (2012b). Talking about how the ways we gain knowledge of our own mental life, Schwitzgebel writes:

> I doubt that we can draw sharp lines through this snarl, cleanly isolating some genuinely introspective process from related, adjoining, and overlapping processes. What we have, or seem to have, is a cognitive confluence of crazy spaghetti, with aspects of self-detection, self-shaping, self-fulfillment, spontaneous expression, priming and association, categorical assumptions, outward perception, memory, inference, hypothesis testing, bodily activity, and who only knows what else, all feeding into our judgments about current states of mind.

> To attempt to isolate a piece of this confluence as *the* introspective process – the one true introspective process, though influenced by, interfered with, supported by, launched or halted by, all the others – is, I suggest, like trying to find the one way in which a person makes her parenting decisions, the one cognitive process behind writing a philosophical essay, or [...] the one cognitive process of taking in a science poster. The causes, the influences, the considerations, are too rich within most cases and too variable between cases for any but a radically pluralist account to do justice to the phenomena. (Schwitzgebel 2012b, 41)

Schwitzgebel offers three kinds of considerations in favour of the moderate (or 'pluralist') picture canvassed in this passage.

First of all, he notices that introspection itself seems to support a pluralist account of self-knowledge: close attention to what one does when one tries to find out how things are with one's mental life reveals no unitary method or process or phenomenology. There is nothing like a specific act of 'looking within' that one always (or most often) performs when one wants to know what one is thinking, desiring, hoping, feeling, etc. In each case, one relies on a plurality of ways of knowing and the ways of knowing one relies on in one case differ from the ways of knowing one relies on in others.

Secondly, given what we know about how the mind works in other cases, a pluralist hypothesis about the nature of self-knowledge is rather plausible. Consider your knowledge of facts concerning your mother. You are not equipped with a single dedicated faculty of *mother*-knowledge. Rather, to find out facts concerning your mother your mind relies on a variety of different processes and methods, recruited opportunistically.[67] Now, even if the case of self-knowledge is different because there are one or more ways of knowing that could only ever

67 Schwitzgebel uses another example: quickly taking in a conference poster does not require a dedicated faculty of *poster-taking-in*.

explain *self*-knowledge, why should any of these ways of knowing be the *only* or *typical* way of gaining self-knowledge? Our mind has many ways of knowing things that are general and versatile rather than applicable only to one type of entity or property – why should they not be routinely used to gain self-knowledge?

Thirdly, the thesis that pluralists deny has received, so far, little or no empirical support. In the last thirty years, the project of isolating processes and mechanisms underpinning other ways of knowing (e.g. vision or memory) has made much progress. The same cannot be said of introspection. On the contrary, while "there are well established partial deficits of memory and vision that suggest a certain degree of functional separability among sub-processes; there is currently no parallel taxonomy of partial introspective deficits – no clear pattern, for example, of functional double dissociations among introspective sub-processes" (Schwitzgebel, ibid.). This provides prima facie (albeit certainly defeasible) evidence in favour of the hypothesis that there is no single way of knowing typical of self-knowledge and picked out by talk of 'introspection'.

These considerations are not decisive or conclusive, of course. But it should be conceded that they cast doubt on *Introspection*, the claim that there is a way of knowing typical of self-knowledge and 'radically' different from the way of knowing typical of knowledge of other minds. Some question that claim by denying that anything like introspection exists. But one doesn't have to take the extreme view that introspection doesn't exist to be sceptical about *Introspection*. One just have to take a pluralist (and, I would say, realistic) stance about how we go about knowing our own mental states.

Let us now move to consider *Inference*. Even if there is no way of knowing typical of self-knowledge, there might be a way of knowing typical of knowledge of other minds – namely, inference from observed behaviour – and this way of knowing might be 'radically' different from whatever plurality of ways of knowing explain self-knowledge. What can be

said in favour or against this claim?

Again, there is an extreme and a moderate reaction. The extreme reaction consists in denying that inference from observed behaviour is a way of knowing other minds. The traditional reason to say this is that inference from observed behaviour is not conducive to knowledge and, therefore, does not explain knowledge. It does not explain how we know *what* other people think or feel. For thoughts and feelings can easily be dissimulated, so behaviour is not a reliable source of information about them. And it does not even explain how we know *that* other people think or feel anything at all. For even if the best explanation of the behaviour I observe in others is the hypothesis that they think or feel something, one shouldn't confuse knowledge with a highly justified conjecture.

Once again, I am prepared to concede that this extreme line is completely misguided. The fact that behaviour is misleading in some cases does not mean that it cannot be reliably trusted in others.[68] And there is no good reason for denying that we can know the truth of a hypothesis on the basis of an inference to the best explanation.[69] Conceding this much, however, leaves room for a more moderate (and more plausible) reaction. The moderate reaction runs as follows. Let it be granted that knowledge of other minds can be gained on the basis of an inference from observed behaviour. Still this doesn't mean that inference from observed behaviour is *the* way of knowing typical of knowledge of other minds. Rather, there may be *many* ways of knowing other minds, none of which is largely or significantly predominant. If this moderate (or 'pluralist') approach is correct, it is simply not true that there is a single way of knowing typical of knowledge of other minds. Epistemologically speaking, knowledge of other minds is no more a unitary category than self-knowledge is: *no* way of knowing (whether inferential or not) is typical of it.

---

68 The point is prominent in McDowell (1982).

69 For a discussion, see Douven (2011).

Given the account of ways of knowing proposed at the beginning of this section, the thesis that no single way of knowing is typical of knowledge of other minds is compatible with several other claims. It is compatible with the claim that there is a certain gerrymandered bunch of activities we engage in whenever we want to know what other people think or feel. And it is also compatible with the claim that it is only to know about others that we engage in that particular gerrymandered bunch of activities – other bunches of activities being required for knowledge in other domains. All the moderate approach says is that there is no single *way of knowing* corresponding to that bunch of activities or, equivalently, that a good explanation of knowledge of other minds will appeal, in each case, to a variety of different methods or processes. In a slogan: knowledge of other minds has no single explanans and, therefore, no dedicated way of knowing.

The moderate approach has, I think, many sympathizers or, at least, many potential sympathizers. One potential sympathizer is Cassam (2007b), according to whom "even if one thinks that knowledge of the existence of other minds *could be* inferential it doesn't follow that our knowledge of their existence *is* primarily inferential" (Cassam 2007b, 159). Our knowledge of other minds can also be perceptual: "there is such a thing as, say, *seeing* that someone else is angry and thereby knowing that he is angry" (Cassam 2007b, 158; my emphasis). To make sense of this proposal, we need to distinguish carefully the claim that we see other people's behaviour (e.g. gestures, facial expressions, etc.) and *infer*, from that behaviour, what they think or feel, from the claim that our visual perceptual states *directly* ascribe mental states to various individuals and, thereby, put us in a position to know that those individuals enjoy those mental states. It is the latter claim that Cassam is interested in defending, at least with respect to a large class of basic and less-than-basic mental states. On this approach, we can see (or otherwise perceive) that a certain person is nervous or calm, agitated or relaxed, happy or unhappy, aggressive or scared, and so on. This is supposed to be

a way of knowing other minds that is both metaphysically and epistemologically direct. It is metaphysically direct because there is a direct causal link between the fact that the person is in a certain mental state and our knowledge that she is in that state. And it is epistemologically direct because, like in other cases of perception, it requires no conscious or unconscious inference from a set of premises to a conclusion.

A perceptual account of this kind is consonant with certain recent hypotheses about the cognitive basis of knowledge of other minds. For example, in the theory put forward by Baron-Cohen (1995), there are four interrelated modules that evolved as a response to the adaptive problem of understanding the minds of other members of our species. One of this modules, called 'intentionality detector', is a perceptual device that takes motion stimuli as inputs and generates contentful states as outputs. Importantly, the function of these contentful states is *not* to attribute behavioural properties that the subject has to 'decipher' or 'interpret' using his or her tacit knowledge of any general psychological theory. What these contentful states do is, rather, representing individuals in the subject's environment as bearers of various basic volitional mental states, in much the same way as visual perception represents objects in the subject's environment as having various basic shapes or colors.

Other psychological theories allow for our knowledge of other minds to be less metaphysically direct, but do not construe it as inferential. According to the so-called 'Simulation Theory', for instance, when we ascribe mental states to someone else, what we do is imaginatively put ourselves in that person's shoes and use our mind to simulate the mental state that that person is in. Predicting other people's decisions is a stock case that simulation theorists aim to model:

> To read the minds of others, [mindreaders] need not consult a special chapter on human psychology, containing a theory about the human decision-making mechanism. Because

> they have one of those mechanisms themselves, they can simply run their mechanism on the pretend input appropriate to the target's initial position. When the mechanism spits out a decisional output, they can use the output to predict the target's decision. In other words, mindreaders use their own minds to ''mirror'' or ''mimic'' the minds of others. (Goldman 2006, 20)

Simulation theorists are keen to emphasize that the 'empathic' processes involved in basic cases of mindreading need *not* be seen as inferential.[70] We don't have to know any neuroscience to find out about of our own mental states. Similarly, we don't have to know any folk-psychology to reliably simulate the mental states that other subjects are in. (It's true that, as theorists, we can divide 'empathic' mindreading processes into a simulation component and a prediction component. But this division is not a sign of epistemological indirectness. The processes whereby the brain keeps track of its own states can also be divided into a detection component and an interpretation component – this hardly makes these processes an indirect way of gaining self-knowledge).

Defenders of the moderate approach do not see the problem of explaining knowledge of other minds as one of *adjudicating* among inferential, perceptual or simulation theories. They reject altogether the thesis that there is a *single* way of knowing in terms of which knowledge of other minds is typically gained. By their lights, different ways of knowing are deployed in each case and the ways of knowing deployed in one case differ from the ways of knowing deployed in others. The idea is that "Mother Nature is a tinkerer, and in designing a skill as complex and as important as mindreading, she has used lots of different tricks. [The result is that] trying to explain everything about mindreading using a single sort of mechanism or process, [...] is a bit like trying to fit a round peg into an irregular trapezoidal hole" (Nichols and Stitch 2003, 101).

---

70 See Goldman (2006, 30-34).

Of course, a pluralist approach of this sort can be proved wrong by empirical research. After all, the question is hardly one that can be settled from the armchair: maybe there is just one correct theory of how we gain knowledge of other minds – it could be one of the theories I've been outlining or a theory that has not been put forward yet. Maybe. But – just as in the case of self-knowledge – multiplicity strikes many as the safer bet. It would be *surprising* if there turned out to be a single way of knowing typical of knowledge of other minds, just as it would be surprising if there turned out to be a single way of knowing facts concerning one's mother: tasks as specific as these do not seem to require a dedicated way of knowing. If this is true, the reasons to question *Inference* go well beyond concerns about the status of inference from observed behaviour as a way of gaining knowledge (as opposed to highly plausible conjectures). They call into question the very project of looking for *the* explanation of knowledge of other minds.

Now, suppose you are convinced by these considerations, as well as by the considerations made earlier against *Introspection*. And suppose you don't want to abandon the mainstream doctrine that no radical metaphysical difference holds between one's own mental lives and the mental lives of others (i.e. *Equality*). Where would this leave you with respect to the task of vindicating *Asymmetry*? You could no longer say that the reason why self-knowledge and knowledge of other minds are 'radically' different is that the way of knowing that typically underpins the first is 'radically' different from the way of knowing that typically underpins the second (for no single way of knowing is typical of self-knowledge and no single way of knowing is typical of knowledge of other minds). Is there something else you could say? As far as I can see, there are at least two alternative routes you might explore, both of which appeal to the notion of a way of knowing. I will conclude this section by saying why I don't find these alternative options especially promising.

One obvious alternative would be to say that self-knowledge and knowledge of other

minds share *no* way of knowing, viz. that no way of knowing can be used to explain both kinds of knowledge. This would make for a pretty radical difference. Moreover, someone who thinks that self-knowledge and knowledge of other minds share no way of knowing is not committed to saying that any single way of knowing is typical of either of them. So this option is compatible with a 'pluralist' approach to self-knowledge and knowledge of other minds.

The main problem with this line is that it does not have much independent plausibility. And it has even less plausibility in the framework of a pluralist account of self-knowledge and knowledge of other minds. If self-knowledge and knowledge of other minds are not unitary epistemological categories (as the pluralist thinks), chances are that at least some ways of knowing will be shared across the board. For example, take Schwitzgebel's (moderate or 'pluralist') account of self-knowledge. On that account, some methods and processes that explain self-knowledge are likely to feature also in the explanation of knowledge of other minds (these include "categorical assumptions, outward perception, memory, inference, hypothesis testing"). Similarly, take Nichols and Stitch's (moderate or 'pluralist') account of knowledge of other minds. On that account, some methods and processes that explain knowledge of other minds are likely to feature also in the explanation of self-knowledge (these include perception and inference from observed behaviour). I'm not saying that Schwitzgebel's version of pluralism about self-knowledge or Nichols and Stitch's version of pluralism about knowledge of other minds are correct, nor that any version of pluralism about self-knowledge or knowledge of other minds that disallows overlap is completely hopeless. I'm just saying that, by committing oneself to the claim that *no* way of knowing is shared across cases of self-knowledge and knowledge of other minds, one commits oneself to a claim that is no less radical (and unlikely to be true) than *Introspection* or *Inference*. Someone who was moved to abandon *Introspection* and *Inference* by their implausibility will hardly find

solace in this option.

The second option is more promising and it consists in saying that even if self-knowledge and knowledge of other minds share some ways of knowing, there is at least one way of knowing they do *not* share. Someone who takes this line could concede that, by and large, the ways of knowing we use to gain self-knowledge and the ways of knowing we use to gain knowledge of other minds are not radically different. At the same time, one could insist that there is at least one way of knowing radically different from any other and unique to one of the two kinds of knowledge. The usual suspect is, of course, introspection, in which case the hypothesis could be the following:

(*Introspection**) There is a radical difference between self-knowledge and knowledge of other minds because there is a way of knowing radically different from any other (namely, introspection) and that way of knowing (unsupplemented with other ways of knowing) could only ever explain self-knowledge.

*Introspection** does not say that introspection is typical of self-knowledge in general – it could equally well be typical of a rather small range of judgments one can make about one's own mental states. The idea is simply that, insofar as one way of knowing is unique to it, self-knowledge remains radically different from knowledge of other minds.

The first thing to be said about this strategy is that its credibility depends largely on its details. One key point that needs to be clarified is the nature of introspection: what is supposed to make it radically different from any other way of knowing? It can't be the fact that it explains knowledge of mental facts as opposed to other sorts of facts, for arguably some ways of knowing associated with knowledge of other minds do that as well (recall Baron-Cohen's 'intentionality detector' and Goldman's simulative mindreading). And it can't

be the fact that the knowledge one gains through introspection is metaphysically or epistemologically direct, for many other ways of knowing have that feature (e.g. perception). Finally, it can't be the fact that it is a way of knowing dedicated to one's own features, for that would effectively reduce *Introspection** to the trivial claim that self-knowledge differs from knowledge of other minds in being knowledge of one's own mind as opposed to the minds of other subjects. What the special nature of introspection might be remains rather obscure.

But there is another, more serious problem with *Introspection**. The problem is that if there was indeed a way of knowing radically different from any other and typical not of self-knowledge in general but of a certain limited range of cases of self-knowledge, we would observe a 'radical' asymmetry not just between self-knowledge and knowledge of other minds, but also between two sorts of self-knowledge: self-knowledge gained through the special way of knowing in question and self-knowledge gained through other, more familiar ways of knowing. But that there is such an 'internal' asymmetry – that is to say, an asymmetry *within* the boundaries of self-knowledge – seems to me highly controversial. In fact, I think there is a strong prima facie case for saying that the very respect in which self-knowledge differs from knowledge of other minds is also one in which self-knowledge presents itself as *homogeneous*: with respect to the ways in which it is gained, self-knowledge is plural and diverse, but with respect to to its 'radical' difference from knowledge of other minds, self-knowledge presents itself as uniform and undifferentiated.[71]

Reflection on this point raises a more general worry about *Introspection**. Let it be granted that introspection exists, is radically different from any other way of knowing and could only ever explain self-knowledge. The question remains whether this is enough to vindicate *Asymmetry* (i.e. the claim that self-knowledge is radically different from knowledge

71 A similar point is made also by Goldman (2006, 227). Notice that the point is not that self-knowledge is homogeneous *in every respect*, but only that it is homogeneous *in the respect in which it is radically different from knowledge of other minds*.

of other minds) or whether it only vindicates the weaker claim that self-knowledge *gained through introspection* is radically different from any other instance of knowledge. Compare: the observation that living in New York is radically different from living in any other place does not vindicate the claim that living in the States is radically different from living in any other country, even if one could only live in the States by living in New York. After all, the observation that living in New York is radically different from living in any other place does nothing to show that living in any region of the States apart from New York is radically different from living anywhere else. Similarly here: the observation that introspection is radically different from any other way of knowing shows nothing about instances of self-knowledge that have *not* been gained through introspection, notwithstanding the assumption that introspection could only ever explain self-knowledge. So even setting aside worries having to do with the homogeneity of self-knowledge, it's not clear to me that a convincing case can be made in favour of *Introspection**.

3. *Asymmetry* and the Subjectivist View of the Mental

MVM is premised on the idea that all subjects are metaphysically 'on a par': no subject is such that his or her mental states are importantly different from the rest. It is because of this commitment to *Equality* that friends of the Mainstream View have been led to explain *Asymmetry* in terms of ways of knowing: if one's mental life is not too different from the mental life of others, it is natural to think that the 'radical' difference between self-knowledge and knowledge of other minds reflects some kind of difference between the *way* one knows one's own mental life and the *way* one knows the mental life of others. In the last section, I raised some doubts about this idea. In this section, I will look at how one can make sense of

Asymmetry if one gives up *Equality* and embraces SVM instead.

According to SVM, there is a radical difference between my mental states and the mental states of all other subjects. The difference consists in the fact that all and only Giovanni's mental states are MENTAL states. In Chapter 1, I tried to convey my intuitive grasp of the notion of a MENTAL state. I said that what Giovanni fears is more quintessentially *fearsome* than what other people fear and what Giovanni desires is more quintessentially *desirable* than what other people desire, that there's nothing as *painful* as Giovanni's pains and nothing as *pleasant* as Giovanni's pleasure.[72] According to SVM, this is because Giovanni's fears, desires, pains and pleasures are (respectively) FEARS, DESIRES, PAINS and PLEASURES. In Chapter 2, the notion of MENTAL state has been put to theoretical use in various ways. I suggested that, among all beliefs, only those which have the simplest and most straightforward correctness conditions are BELIEFS (a BELIEF is correct if and only if its content is true simpliciter).[73] I also suggested, since that every mental state can be given a disjunctive real definition in terms of the corresponding MENTAL state, mental states inherit from MENTAL states not only their correctness conditions (if they have any), but also their causal powers. For example, if my pain causes me to believe *that I am in PAIN*, it does so in virtue of the fact my PAIN causes me to believe *that I am in PAIN*.[74] I take all this to indicate that the distinction between mental states and MENTAL states – besides being grounded in certain deep-rooted intuitions – is also one for which any adequate subjectivist theory will have a lot of explanatory work to do. That distinction can receive new application in the context of a subjectivist account of *Asymmetry*.

---

72 See Chapter 1, § 2.

73 See Chapter 2, § 3.

74 For the principle that every mental state can be given a real definition in terms of the corresponding MENTAL STATE, see Chapter 1, § 3. For the claim that mental states inherit from the corresponding MENTAL STATES their causal powers, see Chapter 2, § 4.

The hypothesis I want to defend is the following: if there's a radical difference between my knowledge of my own mental life and my knowledge of the mental life of others it is simply because there's a radical difference between my mental life and the mental life of others. Differences between ways of knowing, capacities, skills, faculties, methods, processes or cognitive mechanisms at work in the two cases are completely beside the point. The only difference that really matters is intrinsic to the kinds of *facts* I know in the two cases. When I know about my own pain, I know, of a certain individual x, that x has PAIN. When I know about someone else's pain, I know, of a certain individual y, that y has pain. Two different properties, two different facts, hence two different kinds of knowledge.

Let 𝕄-knowledge be knowledge, concerning a certain individual x and a certain MENTAL state 𝕄, that x is in 𝕄 (for example, knowledge *that Giovanni is in PAIN*). And let m-knowledge be knowledge, concerning a certain individual y and a certain mental state M, that y is in M (for example, knowledge *that Fred is in pain*).[75] Then the subjectivist hypothesis about *Asymmetry* can be put, more formally, as follows:

(𝕄-*Knowledge Hypothesis*) There is a radical difference between self-knowledge and knowledge of other minds because cases of self-knowledge are typically cases of 𝕄-knowledge, whereas all cases of knowledge of other minds are cases of m-knowledge.

More crudely put, self-knowledge typically ascribes MENTAL states, whereas knowledge of

75 I use the expression "concerning a certain individual" rather liberally: for a piece of knowledge to concern a certain individual it suffices that it be, in some broad pre-theoretic sense, about that individual. I do not require that a piece of knowledge be *directly referential* for it to qualify as 𝕄-knowledge or m-knowledge: knowledge *that the guy sitting at the counter is thirsty* qualifies as m-knowledge even if the person who knows does not have a thought that directly refers to the guy who is sitting at the counter.

other minds only ascribes mental states.

(Notice that the 𝕄-*Knowledge Hypothesis* is compatible with the existence of cases of self-knowledge that are *not* cases of 𝕄-knowledge. Saying that cases of self-knowledge are *typically* cases of 𝕄-knowledge is not the same as saying that *all* cases of self-knowledge are cases of 𝕄-knowledge (just like saying that dogs typically bark is not the same as saying that *all* dogs bark, to use a familiar example). Are there cases of self-knowledge that are not cases of 𝕄-knowledge? I don't see why not. Right now I know that I am THIRSTY. But being THIRSTY implies being thirsty, because being thirsty is just a matter of being either THIRSTY or otherpersonally THIRSTY. Knowing this much, I also know that I am thirsty. This is a case of self-knowledge: I know myself to be in a certain mental state. It is not a typical or paradigmatic case of self-knowledge, though. If the 𝕄-*Knowledge Hypothesis* is true, self-knowledge is typically 𝕄-knowledge).

Before moving to consider how the 𝕄-*Knowledge Hypothesis* fares as an explanation of *Asymmetry*, let us take a closer look at how exactly it fits with SVM. One important thing to notice is that SVM entails that every case of 𝕄-knowledge is a case of knowledge concerning oneself. The proof begins with a principle that is already part of SVM, namely:

(*Relative Factivity*) For any subject x, if x knows a proposition p then p is true from x's point of view.[76]

𝕄-knowledge is knowledge, concerning a certain individual y and a certain MENTAL state 𝕄, that y is in 𝕄. Given *Relative Factivity*, if a subject x has knowledge, concerning a certain

76 See Chapter 3, § 3.

individual y and a certain MENTAL state 𝕄, that y is in 𝕄, it must be true, from x's point of view, that y in in 𝕄. But SVM says that, for every subject x, from x's point of view only x's mental states are MENTAL states. So if a subject x has knowledge, concerning a certain individual y and a certain MENTAL state 𝕄, that y is in 𝕄, the individual y in question can be no one but x herself. Which is just to say that, if SVM is true, every case of 𝕄-knowledge is a case of knowledge concerning oneself.[77]

The observation is important because it shows that the truth of one of the two components of the 𝕄-*Knowledge Hypothesis* (i.e. the claim that knowledge of other minds can only be m-knowledge) is directly implied by SVM. An objector might point out that SVM does not imply the truth of the other component, i.e. the claim that cases of self-knowledge are typically cases of 𝕄-knowledge (for all SVM says, cases of self-knowledge could be cases of m-knowledge). True. But a subjectivist is under no obligation to show that *every* component of the hypothesis in terms of which she wants to explain *Asymmetry* follows

77 Someone might object that showing that every case of 𝕄-knowledge is a case of knowledge concerning oneself falls well short of showing that every case of 𝕄-knowledge is a case of self-knowledge: self-knowledge is not simply knowledge concerning oneself, but knowledge whereby one thinks of oneself *under a distinctive first-person mode of presentation*. There are two things that a defender of SVM can do in response to this objection. The first is to offer a plausible account of how 𝕄-knowledge can be gained and to show that, on that account, much of the 𝕄-knowledge I can gain is not only knowledge concerning myself, but also knowledge whereby I think of myself under a distinctive first-person mode of presentation. This is what I will do in the next chapter, where I will argue that, most of the time, when I know myself to be THIRSTY, I know that on the basis of the 'Lichtenbergian' judgment *that there is THIRST* – a judgment that I can only translate into non-'Lichtenbergian' terms by using a distinctive first-person mode of presentation. The other thing that a defender of SVM can do is to argue that, once we have a distinction between m-knowledge and 𝕄-knowledge, we can also de-emphasize the idea that self-knowledge involves a distinctively first-person mode of presentation and, perhaps, downplay the philosophical significance of first-person modes of presentation in general. For a deflationary approach to the role of first-person modes of presentation, see Cappelen and Dever (2013).

logically from SVM (compare: an advocate of MVM is under no obligation to show that the hypotheses in terms of which she wants to explain *Asymmetry – Introspection* and *Inference*, for instance – follow logically from MVM). What the subjectivist needs to show is just that every component of her hypothesis, besides being compatible with SVM, fits naturally with it.

This is what we find here. Rather than being an artificial add-on, the claim that cases of self-knowledge are typically cases of 𝕄-knowledge is exactly what one would *expect* given the rest of the subjectivist doctrine. More importantly, it is a claim that, in the context of SVM, can be used to explain a range of different phenomena, and not just *Asymmetry*. For example, consider once again the intuition that one's own mental states are somehow different from the mental states of everyone else. It seems to me that, insofar as they want to endorse that intuition, advocates of SVM should also be able to tell a story about its genesis (if we got it right, it is legitimate to ask *how* we got it right). The 𝕄-*Knowledge Hypothesis* provides the natural basis for that story: the reason why we came to have the intuition that our own mental states are different from the mental states of other subjects is that we typically represent ourselves as having MENTAL states, whereas we always represent others as having mental states.[78] In the next two chapters, I will also show how the idea at the basis of the 𝕄-*Knowledge Hypothesis* – i.e. that knowledge of one's own mental life is involved with MENTAL states – can be used to account for other prima facie puzzling features of our knowledge of the mental – namely the fact that the judgments one forms about what one's own mental life are immune to error through misidentification and the fact that experiential knowledge (i.e.

78 Arguably, the claim that we typically represent ourselves as having MENTAL states implies that we typically deploy the concepts of various MENTAL states. It does *not* imply that all of us possess the general concept of a MENTAL state, nor that we ordinarily make the distinction between mental states and MENTAL states the object of explicit reflection.

knowledge of what it is like to undergo various kinds of experiences) can only be 'first-hand' knowledge. None of this shows that the 𝕄-*Knowledge Hypothesis* is a necessary consequence of SVM. But it seems more than enough to remove worries of artificiality and *ad hoc-ness*.

According to the 𝕄-*Knowledge Hypothesis*, self-knowledge and knowledge of other minds are different because the first is typically 𝕄-knowledge, whereas the second is always m-knowledge. Two questions need to be addressed: is this difference 'radical' enough? And what consequences does the 𝕄-*Knowledge Hypothesis* have for the modal status of *Asymmetry*?

Let me turn to the first question first. A good case can be made that the difference between 𝕄-knowledge and m-knowledge is a pretty 'radical' difference.

First of all, it is not a gradual difference, but a categorical one, because it is a difference in the kind of facts that are known in the two cases and one could not measure or represent such a difference on a scale.

Secondly, if SVM is correct, the difference between 𝕄-knowledge and m-knowledge is a very natural or 'joint-carving' difference – the kind of difference that would figure prominently in any good taxonomy of knowledge. For not only do 𝕄-knowledge and m-knowledge involve the ascription of different properties, but the properties they ascribe are, themselves, deeply different: 𝕄-knowledge ascribes MENTAL states, whereas m-knowledge ascribes states that are *either* MENTAL states *or* otherpersonal MENTAL states (to use a familiar analogy, there is as much difference between the two cases as there is between knowing that there is an explosion and knowing that there is, was or will be an explosion).

Last, but not least, if SVM is true, the kind of difference we observe between 𝕄-knowledge and m-knowledge does *not* hold between different instances of knowledge of

other minds. This is something worth stressing, for a defender of MVM could observe that even on her account self-knowledge and knowledge of other minds target different kinds of facts: when I know about my pain I know the fact *that Giovanni is in pain*, when I know about Fred's pain I know the fact *that Fred is in pain*. But this observation leads nowhere. For of course the same kind of difference that holds between knowing *that Giovanni is in pain* and knowing *that Fred is in pain* holds also between knowing *that Fred is in pain* and knowing *that Ted is in pain*. So if the friend of MVM tries to explain the 'radical' difference between self-knowledge and knowledge of other minds by pointing out that different individuals are involved in the two cases, she will end up with too many 'radical' differences. Not so on the proposal advocated here: the gulf between 𝕄-knowledge and m-knowledge is between myself on the one side and *everyone else* on the other.

The difference between 𝕄-knowledge and m-knowledge is categorical, joint-carving and unparalleled by any other difference that holds between distinct instances of knowledge of other minds. These features make it a good candidate for being the 'radical' difference we are looking for. But is it necessarily true that self-knowledge is typically 𝕄-knowledge and knowledge of other minds is only ever m-knowledge? And if it is not, does this mean that *Asymmetry* could have easily been false?

The answer to these questions depends in part on whether certain claims that SVM makes about subjects and points of view are necessarily true. For example, SVM says that, from any point of view, only *one* subject has MENTAL states. Maybe there are worlds in which this is not so. Maybe there is a world in which, from some point of view, both you and I have MENTAL states. And maybe in such a world I can know you to have MENTAL states and you can know me to have MENTAL states – a counterexample to the claim that knowledge of other minds is invariably m-knowledge. Or take another example. SVM says that every subject gets only one subjective chance to be the special one, i.e. that for every subject there is just one

point of view from which that subject has MENTAL states. Again, there might be worlds in which this is not so. Maybe there is a world in which some of my mental states are MENTAL states, but many others are only MENTAL states from some other point of view. And maybe in such a world I mostly know myself to have otherpersonal MENTAL states – a counterexample to the claim that in all possible worlds self-knowledge is typically 𝕄-knowledge.

Maybe. Or maybe not. My inclination is to think that there can't be points of view from which more than one subject is special and that there can't be subjects that get more than one subjective chance to be special. I take it to be in the nature of subjects and points of view that they have to be related in these ways (just as I take it to be in the nature of times and instants that there can't be times at which more than one instant is present and that there can't be instants that get more than one fleeting occasion to be present). But suppose I am wrong. Suppose there are possible worlds in which the relation between subjects and points of view is not as harmonious as it is in the actual world. A plausible case can still be made that these possible worlds (if they exist) are extremely remote from the actual one. After all, a world in which numerically distinct subjects have MENTAL states would be entirely *unlike* the world as I found it (it would certainly be more dissimilar from actuality than the world described by McGinn (2004), where subjects have been wired up in such a way that they have indirect knowledge of their own mind and direct knowledge of other minds). So even if the 𝕄-*Knowledge Hypothesis* gives us only a contingent version of *Asymmetry*, this doesn't imply that *Asymmetry* could have *easily* been false.

4. *Asymmetry* and the Privacy of the Mental

SVM says that, for each subject x, there is one and only one point of view from which x's mental states are MENTAL states. As we've seen in the last section, this is part of the reason why, if SVM is correct, every instance of 𝕄-knowledge is an instance of knowledge concerning oneself. But the principle has another, more dramatic consequence, namely that every instance of 𝕄-knowledge is an instance of *private* knowledge: when I have knowledge, concerning a certain individual x and a certain MENTAL state 𝕄, that x is in 𝕄, nobody else can know what I know. Since whatever I know has to be the case from my point of view and my point of view is the point of view from which all and only *my* mental states are MENTAL states, I can only know *myself* to have MENTAL states. As to other people, whatever *they* know has to be the case from *their* point of view and their point of view is one from which all and only *their* mental states are MENTAL states, so they can only know *themselves* to have MENTAL states. There's no chance that I can know *you* to have MENTAL states or that you can know *me* to have MENTAL states.

I regard this consequence as an alluring feature of the account of *Asymmetry* defended here. That the knowledge I have of my own mental life is in some sense *private* and that this privacy is closely connected with the asymmetry between self-knowledge and knowledge of other minds is, I submit, as deep-seated an intuition as *Asymmetry* itself. Sometimes, this is put by saying that our mental life is 'hidden' from other subjects:

> Man, though he have great variety of thoughts, and such from which others as well as himself might receive profit and delight; yet they are all within his own breast, invisible and hidden from others, nor can of themselves be made to appear. (Locke 1975, 404-405)

Other times, the metaphor is that our life is 'inaccessible' to others:

> One of the characteristics which are ascribed to mental objects or events is that they are in some way private. Thus it is commonly held that out thoughts and feelings, our dreams and imaginings, our sensations and memories, are things to which we alone have access. (Ayer, 1990, 199)

Of course, if you're an advocate of MVM, you will refrain from fully endorsing these intuitions. You will not say that the mind of another is something one is unable to inspect. Rather you might say that "the mind of another is something which one is unable *directly* to inspect" (Lewis 1964, 332). Nor will you say that one person has no access of any sort to the events of the inner life of another. You might say, instead, that "one person has no *direct* access of any sort to the events of the inner life of another" (Ryle 1949, 13). It's clear that these reformulations – which rest on a (controversial) distinction between direct and indirect ways of knowing – do not make the mental truly private. Only subjectivists seem to be in a position to capture the idea that, when I know about my mental life, I knows some facts (e.g. this pain being PAINFUL, this belief being a BELIEF, etc.) that nobody else can know.

I said that this is an alluring feature of the view, but others will describe it as a fatal flaw. Subjectivists posits subjective facts alongside objective ones. But then they go on to say that these subjective facts, or a large portion of these subjective facts, cannot be known by anyone but the subject from whose point of view they obtain. It might be complained that this amounts to an unacceptable form of epistemological solipsism. Just like certain sceptics think that we can know little or nothing about other people except facts concerning their behaviour, subjectivists seem to think that we can know little or nothing about other minds except facts involving ordinary ('lowercase') mental states. You can know my headache to be painful, but you cannot know it to be PAINFUL. Nor can you know any of my beliefs, hopes, desires,

feelings and experiences to be BELIEFS, HOPES, DESIRES, FEELINGS and EXPERIENCES: the glowing side of my mental life is something of which I am destined to be the only witness. Some will regard this result as implausible or undesirable. Others might take it to be a refutation of SVM: if so many subjective facts turn out to be 'private', the question arises whether they are genuine facts (rather than, say, subject-relative representations of a perfectly objective reality).

There are various things that subjectivists can and should say in response to these accusations. First of all, they should defend the coherence of the notion of a private fact – i.e. a fact that can only be known (if it can be known at all) by the subject from whose perspective it obtains. They should point out that, metaphysically speaking, private facts are facts like all others. In effect, a private fact is nothing more than the subjective analogue of an instantaneous fact – i.e. a fact that can only be known (if it can be known at all) at the unique time at which it obtains. If there are temporary facts (as many philosophers believe), chances are that there are also some instantaneous facts. For example, *that this instant is present* is very plausibly an instantaneous fact: there is no question of knowing that fact at any time other than this time. There seems to be nothing wrong with instantaneous facts – in particular, it's difficult to see why their being instantaneous should be incompatible with their being genuine facts. Similarly, there seems to be nothing wrong with private facts – just because they are private this doesn't mean that they don't qualify as facts or that they cannot do all the things that other facts can do (e.g. make propositions true or false, cause other facts, distinguish the actual world from various counterfactual possibilities, etc.).

Secondly, subjectivists should emphasize the distinction between epistemological solipsism – the claim that I can know *nothing* about your mental life and you can know *nothing* about mine – and a moderate privacy thesis – the thesis that there is *something* you know about your mental life that I cannot know and there is *something* about my mental life

that you cannot know. There is little reason to think that SVM should imply anything like epistemological solipsism. On the other hand, the moderate privacy thesis is one that, as already noted, has a great deal of prima facie intuitiveness – in fact, subjectivists should present it as a profession of sensible epistemological modesty.

There's also another distinction that subjectivists should draw in this connection. The distinction I have in mind is between the following two theses:

(1) Otherpersonally, there are facts about your mind that I cannot know

(2) There are facts about your mind that I cannot know.

It's true that SVM implies (1), but subjectivists can and should resist the inference from (1) to (2), just as temporaneists can and should resist the inference from (1') to (2'):

(1') In the future, there will be facts that, at present, cannot be known

(2') There are facts about the future that, at present, cannot be known

In a temporaneist framework, one cannot go from the future existence of facts that are presently unknown to the present existence of unknown facts about the future. Similarly, in a subjectivist framework, one cannot go from the otherpersonal existence of facts that are firstpersonally unknown the firstpersonal existence of unknown facts. It seems to me that, in the light of these distinctions, the moderate privacy thesis implied by SVM gains even further plausibility: subjectivists do not ask us to believe in the existence of facts that form part of reality from every subject's point of view, but are mysteriously 'hidden from sight' from some of them; they ask us to believe in the existence of facts that are 'hidden from sight' from

precisely those points of view from which they do not obtain.

Finally, subjectivists should be careful to make clear that, even if I cannot know your beliefs, hopes, desires, feelings and experiences to be BELIEFS, HOPES, DESIRES, FEELINGS and EXPERIENCES, this doesn't mean that I can only know you to have ordinary ('lowercase') mental states. There is something else I can know, namely that your beliefs, hopes, desires, feelings and experiences are *otherpersonally* BELIEFS, HOPES, DESIRES, FEELINGS and EXPERIENCES. In other words, I cannot be witness of the glowing side of your mental life, but I can know that, *from your point of view*, your mental life has the same glowing side my mental life has from mine.

CHAPTER 4:

Immunity to Error Through Misidentification

When a judgment goes wrong, it can go wrong in many different ways. For instance, suppose I am in a bar, making various judgments about what is happening around me. There are many kinds of mistakes I can make. I may judge *that there is someone sitting at the counter* when, in fact, no one is sitting at the counter. I may judge *that the person sitting at the counter wears blue trousers*, when, in fact, the person sitting at the counter wears black trousers. Or I may judge that *Paul McCartney is sitting at the counter*, when, in fact, someone is sitting at the counter, but he is not Paul McCartney. Errors of the last sort – whereby one mistakes someone for someone else – are often referred to as cases of *error through misidentification*.

It was Wittgenstein who first pointed out, in a famous passage of the *Blue Book*, that in the case of the judgments one forms about one's own mental life the possibility of an error through misidentification "has not been provided for" (1958, 66-67). For instance, suppose I'm sitting in a bar, waiting for my drink, and think to myself *that I'm thirsty*. Judgments of this sort can certainly go wrong: sometimes one takes oneself to be thirsty even when one is not thirsty. But they do not seem to be the kind of judgments that can go wrong through misidentification, viz. as a result of mistaking someone who is thirsty for one's non-thirsty self. I will follow the tradition in calling this phenomenon *immunity to error through misidentification*:

(*Immunity*) The judgments I form about what my mental life are immune to error through

misidentification.

The exact nature and scope of this phenomenon are a matter of controversy and I will spend a great amount of time clarifying its contours. What should be clear already is why – since Wittgenstein's remark was first published and Shoemaker (1968) brought the point to everyone's attention – immunity to error through misidentification has caused the philosophical concern it has. The problem is that it is not clear why the judgments I form about my mental life should display any kind of immunity to error through misidentification if it is true that:

(*Particularity*) The judgments I form about my mental life are judgments about a particular individual, namely myself.

Given *Particularity*, one would expect the judgments I form about my mental life to be just as likely to go wrong through misidentification as other types of particular judgments. After all, if a particular individual (namely, myself) is at stake when I form these judgments, shouldn't it be routinely possible to mistake someone else for that individual, just as it is routinely possible to mistake someone else for Paul McCartney? And yet this is not what we observe. What we observe is that the judgments one forms about what one believes, hopes, desires, feels, experiences, etc. are (in some yet to be fully clarified sense) immune to error through misidentification. Hence the philosophical discomfort: given *Particularity*, *Immunity* stands in need of explanation.

The problem has, I think, a direct relevance for the dispute between the Mainstream View of the Mental (MVM) and the Subjectivist View of the Mental (SVM). I will argue that advocates of MVM have no plausible way of denying *Particularity*. This leaves them with

little choice but to search for some kind of epistemological explanation of *Immunity* – a search that has so far been unsuccessful (§ 2). On the other hand, we will see that if one embraces SVM, one can reject *Particularity* and offer a distinctively metaphysical explanation of why *Immunity* holds (§ 3). This chapter will outline, discuss and compare these two strategies. But in order to do that, it is indispensable to get clear what we talk about when we talk about immunity to error through misidentification. A definition of error through misidentification has to be offered and different senses of immunity have to be distinguished. This is the task of the next section.

1. Immunity

What is an error through misidentification? What does immunity to such errors consist in? And what exactly is involved in the task of explaining immunity to error through misidentification? Let me address each of these questions in turn.

1.1 What is an error through misidentification?

I take error through misidentification to be a feature of certain thoughts or judgments (I will use the two terms interchangeably). In looking for a definition of *error through misidentification* (EM, hereafter), what we are looking for is a set of (non-trivial) necessary and sufficient conditions for a thought or judgment to qualify as a case of error through misidentification. The project is subject to two kinds of criticisms.

One criticism is that the notion of EM is a technical notion and we do not *look* for definitions of technical notions, we *stipulate* them. So we shouldn't look for a definition of EM, but simply stipulate what we want the expression "error through misidentification" to mean.

This criticism strikes me as misguided. The label of "error through misidentification" may sound technical, but the notion of an error through misidentification is not. An error through misidentification occurs when, in making a judgment or forming a thought that is otherwise correct, someone "misidentifies" something with something else or (perhaps more neutrally) "mistakes" something for something else. To illustrate, consider again the example we started with:

> *Paul McCartney* – I'm in a bar and look at the man who's having a beer at the counter. The man looks impressively like Paul McCartney, so I think to myself that *Paul McCartney is sitting at the counter*. Then I get closer and realize that I was wrong: someone is sitting at the counter, but he is not Paul McCartney.

In this case, one would ordinarily say that I "misidentified" the guy who is sitting at the counter with Paul McCartney, or that I "mistook" him for Paul McCartney. Our intuitive grip on the use of these expressions is reasonably good. At least, it is as good as our intuitive grip on the use of other locutions that are of interest to epistemologists, such as "is justified to believe" or "is evidence for". In looking for formal definitions, it is this intuitive grip that we are trying to clarify and make more precise.

A rather different criticism is that, precisely because the notion of EM is an ordinary notion, it is hopeless to try to analyse it in terms of a set of (non-trivial) necessary and sufficient conditions. After all, successful examples of conceptual analysis are the exception, not the rule, when ordinary concepts are at issue.

While I am sympathetic to its spirit, I don't think that this criticism completely undermines my project. As Williamson observes with respect to the task of defining the notion of *evidence*, "indiscriminate description of the ordinary use of a term and arbitrary stipulation of a new use are not the only options. We can single out theoretical functions central to [a certain ordinary concept], and ask what serves them" (Williamson 2000, 194). This is what I will try to do here: different candidate definitions of EM will be tested against our intuitions about a range of cases, but also assessed on the basis of their theoretical virtues. My guiding principle will be that a good definition of EM should satisfy two criteria: it should be *specific* enough to capture what is distinctive of the mistake I commit when I judge *that Paul McCartney is sitting at the counter* while someone else is, but also *broad* enough to encompass the full range of judgments whose 'epistemic badness', from a purely theoretical perspective, is not too dissimilar from that of my judgment about Paul McCartney.

With these criteria in mind, let us look at some recent attempts to define EM. A good starting point for our discussion is the definition of EM offered by François Recanati in his recent "Immunity to Error through Misidentification: What It Is and Where It Comes From" (2012b). Recanati suggests that EM occurs when "a subject S judges that some object *a* is F, because S has grounds for believing that some object is F and wrongly believes that *a* is one such object", viz. "an object satisfying (λz) (S has grounds for believing that *z* is F)" (Recanati 2012b, 180). More schematically:

(Recanati) A judgment J is an error through misidentification if and only if:

(i) J is a singular judgment that some object a is F;

(ii) the author of J has grounds for believing that some object is F;

(iii) the author of J wrongly believes that she has grounds for believing that a is F.

Here's what Recanati would say about the Paul McCartney case: my judgment is a singular judgment that a certain individual, Paul McCartney, is sitting at the counter, I have grounds for believing that someone is sitting at the counter, but I am wrong in believing that I have good grounds for thinking that it is Paul McCartney who is sitting at the counter. Taken together, these features make my judgment a case of error through misidentification.

The main problem with this proposal has to do with condition (iii). When, in the situation envisaged above, I judge *that Paul McCartney is sitting at the counter*, it's not clear that I have to have or form any second-order belief about my grounds for believing that Paul McCartney is sitting at the counter. In particular, it's not clear that I have to have or form a second-order belief to the effect that I have good grounds for believing that Paul McCartney is sitting at the counter: I may simply form a judgment about Paul McCartney. What's more, even if I were to form some second-order belief to the effect that I have good grounds for believing that Paul McCartney is sitting at the counter, it's not clear that my second-order belief would be *wrong*. After all, I have plenty of good grounds for believing that Paul McCartney is sitting at the counter: I'm looking at the guy over there, he is sitting at the counter and he looks impressively like Paul McCartney.

One may suggest that the mistake in Recanati's definition is the following: what's distinctive of EM is *not* the fact the author of the judgment wrongly believes herself to have grounds for the judgment, but, more simply, the fact that the grounds she has for the judgment are somehow misleading. Plausibly, in the Paul McCartney example I have evidence or grounds for making my judgment, but the evidence or grounds in question are grounds for believing the false proposition *that the guy I'm looking at is Paul McCartney*, or, more simply, *that that guy is Paul McCartney*. Building on this suggestion, one might be led to define EM in terms of Pryor's (1999) notion of 'de re misidentification':

(Pryor-1) A judgment J is a case of *de re* misidentification if and only if:

(a) J is a singular judgment that some object a is F;

(b) the author's grounds for J are grounds for believing that some object b is F and that b is identical to a;

(c) however, unbeknownst to the subject, a and b are not identical.[79]

This definition applies smoothly to the Paul McCartney example we started with: contrary to my assumption *that that guy is Paul McCartney*, that guy is not Paul McCartney and this seems to capture the sense in which my judgment, being based on grounds that justify that identity assumption, involves an error through misidentification.

As Pryor himself noticed, however, some cases make trouble for *de re* misidentification:

> *The skunk* – I smell a skunky odor and see several animals rummaging around in my garden. Approaching closer and sniffing, I form the belief, of the smallest of these animals, *that it is a skunk in my garden*. The belief is mistaken. There are several skunks in my garden, but none of them is the small animal I see.

Consider my judgment or belief, of the smallest of the animals rummaging in my garden, that it is a skunk. The evidence I have for making that judgment does not seem to include or justify any identification of the form 'b is identical to a', contrary to what condition (b) demands: for my judgment to be justified, there need not be any particular skunk such that I can reasonably believe, of the smallest of the animals rummaging around in my garden, that it

79 See Pryor (1999, 274-275).

is identical with *that skunk*.[80] Despite this, it seems natural to say that my judgment involves some species of EM: after all, the judgment is wrong and it is wrong because I mistook the smallest of the animals in my garden for one of the skunks in my garden. Based on this kind of considerations, Pryor argues that not all instances of EM are also cases of *de re* misidentification and, along with many others, I agree with him on this point.

However, Pryor suggests that *de re* misidentification constitutes an interesting *subspecies* of EM and this is where I part company with him. The crucial point is that for a judgment *that a is F* to qualify as a case of *de re* misidentification there doesn't even have to be *something* which is F. Hence cases like the following count as cases of *de re* misidentification:

> *The Lion King* – An otherwise very trustworthy person tells me that Al Pacino and Harrison Ford are the same person. She also tells me that Harrison Ford starred in "The Lion King". On this testimonial basis, I come to justifiedly (and yet wrongly) believe *that Al Pacino starred in "The Lion King"*.

My judgment *that Al Pacino starred in "The Lion King"* satisfies all the criteria for being a case of *de re* misidentification (it is singular and the grounds on which it is based are grounds that justify a false identification). But I'm inclined to think that it doesn't involve any genuine EM. In any standard case of EM, the author of the judgment, while wrong about the fact that such-and-such person is so-and-so, is at least right about the general fact that *someone* is so-and-so – that's why we say that EM occurs when someone is only wrong about *which* person or thing is so-and-so. Now, since "The Lion King" is an animation movie, a good case can be made that *nobody* starred in it. This means that my judgment *that Al Pacino starred in "The*

80 In fact, Pryor makes the stronger point that I might not yet be in a position to hold any *de re* beliefs whatsoever about any of the things I'm smelling (1999, 282).

Lion King" cannot be described as the result of mistaking the person who starred in "The Lion King" for Al Pacino: there is no such person. But if there is no such person, there is no real *misidentification* going on – so why talk of error *through misidentification*? Sure enough there is a false identification in the background (the proposition *that Al Pacino is Harrison Ford*). But the label "error through mis*identification*" should not mislead us into thinking that the presence of a false identification in the background of a judgment is, all by itself, sufficient to make it a case of EM. Committing what Pryor calls 'de re misidentification' is, in my opinion, neither necessary *nor sufficient* for committing an EM.

Pryor defines another notion of misidentification, which doesn't suffer from the same difficulties. He calls this other notion 'which-object misidentification' (or 'wh-misidentification' for brevity):

(Pryor-2) A judgment J is a case of wh-misidentification if and only if:

(I) J is a singular judgment that some object a is F;

(II) The evidence supporting J puts the subject in a position to know that something is F;

(III) J is false.[81]

Prima facie, the notion of wh-misidentification is better-behaved than the notion of *de re* misidentification. For example, my judgment in *The Skunk* doesn't count as a case of *de re* misidentification, but it does count as a case of wh-misidentification: it is a judgment, of a certain animal, that it is a skunk in my garden and, while it is a false judgment, the evidence supporting it puts me in a position to know that there is some skunk in my garden. On the

81 See Pryor (1999, 282). Pryor suggests a 'small amendment' to condition (II), but the amendment won't matter too much for my present purposes.

other hand, my judgment in *The Lion King* (which counted as a case of *de re* misidentification) does not count as a case of wh-misidentification, because the evidence on which my judgment *that Al Pacino starred in "The Lion King"* is based does *not* put me in a position to know that *someone* starred in "The Lion King". All this is just as it should be.

But there are problems with the notion of wh-misidentification, too. First of all, this notion fails, once again, to cover all the intuitive cases of EM. Consider the following example:

> *The red shirt* - Sitting at the restaurant, looking at the man at the counter, I think to myself "Paul McCartney is wearing a red shirt". As a matter of fact, the man I'm looking at is not Paul McCartney. As it happens, Paul McCartney *is* wearing a red shirt at the time of my judgment.

Notice that my judgment *that Paul McCartney is wearing a red shirt*, being true, involves no wh-misidentification. Now, this is fair enough as far as it goes (after all, one might insist that what is true shouldn't be said to involve any "error" through misidentification). Still, it's clear that the 'epistemic badness' displayed by my judgment *that Paul McCartney is wearing a red shirt* makes it pretty much *akin* to my judgment in *Paul McCartney.* And it is also clear that, to the extent that my judgment *that I'm thirsty* is immune to error through misidentification, it is also immune to the 'epistemic badness' in question. This suggests that, unlike the notion of wh-misidentification, a good working notion of EM should be broad enough to encompass my judgment in *The Red Shirt.*

But the problems don't end here, because, in other respects, the notion of wh-misidentification is also too broad. To see this, reflect on the following example:

> *The oak-tree* - Walking in a forest, I come across two trees, A and B. My body of evidence

> includes two propositions: I see *that B is an oak-tree* and I notice *that A looks pretty much like B*. Based on such evidence, I conclude *that A is an oak-tree*. But A is not an oak-tree.

I think you'll agree that my judgment *that A is an oak-tree*, though wrong, does not involve any EM. It's true that I arrived at the wrong judgment by comparing A to B, but it is not as if I mistook A for B or viceversa. Notice, however, that both the proposition *that B is an oak-tree* and the proposition *that A looks pretty much like B* form part of the evidence I have for believing that A is an oak-tree. Given that I know both of these propositions (and, in particular, the first), it is therefore true, in this case, that the evidence I have for the judgment *that A is an oak-tree* puts me in a position to know the general fact *that something is an oak-tree*. So my judgment that A is an oak-tree satisfies all the requirements for being a case of wh-misidentification.

Where do we go from here? What I want to do is suggest two amendments to Pryor's notion of wh-misidentifcation. The first amendment concerns condition (III). Cases like *The Red Shirt* can be taken to show that a judgment need not be false to exhibit the 'syndrome' of misidentification. But if it is not falsity, what is the bad-making feature of the syndrome? One might be tempted to say that the relevant judgments, though perhaps true, are not justified. But, in any typical case of EM, the subject *is* justified in forming the judgment he or she forms (this was one of the chief reason to reject Recanati's definition of EM). So the bad-making feature is not lack of justification either. One natural possibility remains: the landmark of EM is not lack of truth or lack of justification, but rather lack of knowledge. My judgment *that Paul McCartney is wearing a red shirt* may well be true, but it doesn't amount to knowledge: it's not Paul McCartney that I'm looking at, so the perceptual evidence I have does not put me in a position to know that Paul McCartney is wearing a red shirt (and this is so even if, as it happens, it is true that Paul McCartney is wearing a red shirt). In this respect, my judgment in *The Red Shirt* deserves to be put on the same level as my judgment *that Paul*

McCartney is crossing the street: both are cases in which I fail to know the content of my judgments. So the first suggestion I want to put forward is the following: instead of requiring that a judgment be false in order to count as a case of EM, we should just require that it fail to constitute knowledge.

The second amendment concerns condition (II). Cases like *The Oak Tree* argue for a qualified reading of this condition. There's certainly a viable conception of evidence on which both my belief *that B is an oak-tree* and my belief *that A looks pretty much like B* form part of the evidence I have for believing that A is an oak-tree. However, there seems to be also a notion of *minimal evidence* on which neither of these two beliefs enter into the minimal evidence I have for my (wrong) singular judgment about A. If we think of evidence in terms of propositions and think of proposition in terms of sets of possible worlds, we can think of the minimal evidence a subject has for making a certain judgment as the largest set of possible worlds that justifies the subject in making that judgment. If we conceive of propositions as more fine-grained than sets of possible worlds, we could think of the minimal evidence for a judgment as the weakest proposition that provides the subject with good grounds for that judgment (where a proposition is weaker than another if it analytically entails it). I offer these as possible ways of glossing a notion that I am happy to treat as a primitive and on which we have – I think – a reasonably good intuitive grip. For example, I take it that, intuitively, the minimal evidence I have for thinking that A is an oak-tree doesn't involve anything about B. Rather, it reduces to something along the following lines: *A looks pretty much like what an oak-tree would look like (if there were any oak trees)* or, even better, *A looks oak-tree-ish*. Now, clearly the mere observation *that A looks oak-tree-ish* does not justify me in believing that there are any oak-tree candidates other than A.[82] So, since A is *not* an oak-tree, the

82 Perhaps, if certain versions of externalism about mental content are true, the observation *that I possess the concept of something looking '*oak-*tree-ish'* would justify me in believing that there actually are some *oak*-trees. But of course the observation *that something looks oak-treeish* does not justify me in believing *that I*

minimal evidence I have for believing that A is an oak-tree does *not* put me in a position to know that something is an oak-tree. And this seems to be why *The oak-tree* is *not* a case of error through misidentification. So my second suggestion is that, for a judgment to count as a case of EM, we should require that the *minimal evidence* (and not just any kind of evidence) supporting the judgment put its author in a position to know the relevant existential generalization.

If we implement the two amendments I've suggested, the definition of EM we arrive at is the following:

> A judgment J is an error through misidentification iff:
> - J is a singular judgment that some object a is F;
> - the minimal evidence supporting J puts the subject in a position to know that something is F (hereafter: the minimal evidence supporting J *offers* ∃-*knowledge*).
> - J fails to constitute knowledge;

This definition classifies correctly all the cases discussed in this section, striking a good balance between specificity and broadness.[83] Needless to say, those who feel nervous about calling 'errors' judgments that are perfectly correct (e.g. my judgment in *The Red Shirt*) can

---

possess the concept of something looking oak-*tree-ish*. So the observation *that something looks oak-tree-ish* does not justify me in believing that there actually are oak-trees, even if externalism about mental content is true.

83 Notice that the definition I'm proposing (like the other definitions of EM examined in this section) implies that only *singular de re* judgments can be cases of EM. This requirement strikes me as relatively innocuous, given that what we are ultimately interested in is the immunity to EM of mental self-ascriptions and mental self-ascriptions are singular first-person judgments. At any rate, it's easy to see how the definition could be modified so as to allow for cases of EM that are neither singular nor *de re*.

reject my amendment to condition (III) and work with a more narrow notion of error through misidentification *strictu sensu*:

> A judgment J is an error through misidentification *strictu sensu* iff:
> - J is a singular judgment that some object a is F;
> - the minimal evidence supporting J offers ∃-knowledge;
> - J is false;

Alternatively, one can reserve the label of 'error through misidentification' for the notion just defined, and describe the type of case defined above as one of 'ignorance through misidentification'. Once the distinction is clear, the terminology doesn't matter too much.

1.2. What is immunity to error through misidentification?

As we've seen at the beginning, it was Wittgenstein who first pointed out that, in the case of mental self-ascriptions, “no error [through misidentification] is possible” (Wittgenstein 1958, 67). And it is first and foremost the *impossibility* of any misidentification that Shoemaker (1968) had in mind when he invented the label of 'immunity to error through misidentification'.[84] Now that we know what an error through misidentification is (or at least have a credible hypothesis as to what it is), we should try to get clearer on the nature of this impossibility.

84 “The statement 'I feel pain' [is immune] to error through misidentification [...]: *it cannot happen* that I am mistaken in saying 'I feel pain' because, although I do know of someone that feels pain, I am mistaken in thinking that person to be myself” (Shoemaker 1968, 557).

Let us begin by reminding ourselves that, like all other modal claims, impossibility claims can be *de dicto* or *de re*. It is one thing to say that it is impossible that a bachelor be married (*de dicto* impossibility) and it is another thing to say, of a particular bachelor, that it is impossible that he be a married (*de re* impossibility). So consider my judgment *that I'm thirsty* and consider Wittgenstein's claim that, in this and other cases, "no error [through misidentification] is possible". Is this supposed to imply that my judgment couldn't have been an error through misidentification (*de re* impossibility)? I suppose not. My judgment *that I'm thirsty* can be seen as a particular mental episode or event, but when we consider that judgment in the context of discussing IEM we are not interested in what could (or could not) have been true of that mental episode or event in various counterfactual circumstances. In particular, we are not interested in whether or not that mental episode or event could have been a case of error through misidentification had it had completely different features than the ones it actually has.[85] Presumably, then, the kind of impossibility that underlies IEM is some kind of *de dicto* impossibility. More precisely, the point must be that there's some kind Φ that my judgment *that I'm thirsty* actually exemplifies such that:

(*Impossibility*) Necessarily, for any judgment J, if J is of kind Φ, then J is not a case of error through misidentification.

Notice that *Impossibility* is formulated as a schematic principle: 'Φ' is a schematic letter that

85 An obvious reply is that that it shouldn't be assumed without argument that my judgment could have had completely different features than the ones it actually has. But, even so, the point remains that questions of *de re* modality seem completely off topic in this context. One could be an *haecceitist* (i.e. embrace the view that any entity can possess radically different properties in different possible worlds) and yet recognize a sense in which my judgment *that I'm thirsty* is immune to EM. Symmetrically, one can be an *essentialist* (i.e. embrace the view that every feature of any entity is essential to it) and yet recognize that certain judgments which do not happen to involve any EM are nonetheless not *immune* to EM.

needs to be replaced by a kind term to obtain a true (or false) claim. In this way, no assumption is made about the nature of the kind in question, leaving room for different hypotheses for why, in the case of my judgment *that I'm thirsty*, "no error [through misidentification] is possible". This observation will become especially relevant in the next section.

Now, I said that when Shoemaker invented the label of 'immunity to error through misidentification' what he had in mind was, first and foremost, the impossibility of a misidentification. But if it is to be a distinctive feature of certain selected categories of judgments (like, for instance, the judgments I form about my mental life), immunity to error through misidentification cannot consist merely in the impossibility of a misidentification. To see why, consider ERROR-FREE, the kind including all and only the judgments that do not involve any error through misidentification:

x belongs ERROR-FREE iff x is a judgment that does not involve any EM

Notice that ERROR-FREE satisfies *Impossibility*: necessarily, if a judgment belongs to ERROR-FREE, it does not involve any EM. But is this a good reason to speak of 'immunity to error through misidentification'? It seems not. For example, suppose that, seeing my friend Fred walking along the river, I judge *that Fred is walking along the river*. Even if my judgment is correct and belongs to ERROR-FREE, one wouldn't say that it is immune to error through misidentification in the same sense in which the judgments I form about my mental life are. In fact, my judgment *that Fred is walking along the river* does not seem to be immune to error through misidentification in any philosophically interesting sense (and this is not just because the judgment, regarded as a particular mental episode or event, belongs to ERROR-FREE only contingently – presumably, we wouldn't describe it as "immune" to error

through misidentification, even if we believed that all the properties of a judgment, including its being error-free or not, are essential to it).

Here's another illustration of the same point. Consider KNOWLEDGE, a kind defined as follows:

x belongs to KNOWLEDGE iff x is a judgment that amounts to knowledge

Again, KNOWLEDGE satisfies *Impossibility*: necessarily, if a judgment belongs to KNOWLEDGE, it does not involve any EM. But we would not say that, whenever a judgment constitutes knowledge, it is immune to error through misidentification in the same sense in which the judgments I form about my mental life are. In fact, many judgments that constitute knowledge do not seem to be immune to error through misidentification in any philosophically interesting sense (and this is not just because they belong to KNOWLEDGE only contingently – presumably, we wouldn't describe then as "immune" to error through misidentification, even if we believed that all the properties of a judgment, including whether or not it amounts to knowledge, are essential to it).

These examples seem to me to point in the same direction: for there to be IEM more is needed than just the satisfaction of *Impossibility*. What is the additional ingredient? The suggestion I want to put forward is inspired by the following passage from Campbell (1999):

> A judgment like "Bill spoke", when made on an ordinary basis, is subject to error through misidentification, in that *you could have a ground for doubt about the correctness of the judgment* which did not undermine your right to claim to know, on that basis, the existential proposition, "Someone spoke". (Campbell 1999, 89; my emphasis)

Part of what Campbell says in this passage is that, in the case of many ordinary judgments, the author of the judgment cannot rule out the possibility of his or her judgment being a case of misidentification. The implicit suggestion is that this does *not* happen in any paradigmatic case of immunity: where there is immunity to error through misidentification *the author of the judgment* (and not just any well-informed onlooker) need not worry about the possibility that his or her judgment involve any error through misidentification. Phenomenologically, the suggestion strikes me as rather plausible: in judging *that I'm thirsty*, I am not completely sure to be right (after all, sometimes one thinks one is thirsty even when one is not thirsty), but something about my judgment that I am not committing an error *through misidentification* (it would make hardly any sense *for me* to wonder whether I'm mistaking someone else's thirst for mine). The idea, then, is that IEM doesn't consist just in the impossibility that an error through misidentification might occur, but involves also some sort of *epistemic assurance*, on the part of the author of the judgment, against such a type of error.

What does this epistemic assurance consist in exactly? As a first stab, one might try to define it as follows:

(*Assurance*-1) Necessarily, for any judgment J, if J is of kind Φ, then the author of J is in a position to know, while he or she makes J, that, in so doing, he or she is not committing an error through misidentification.

But there are at least two problems with *Assurance*-1. The first and most obvious is that one should not expect every author of a paradigmatic error-free judgment to possess the concept of an EM (let alone to have all the right philosophical beliefs about EM). This means that – on a fairly natural understanding of 'being in a position to know' – the author of a paradigmatic error-free judgment might *not* be in a position to know that her judgment is not a case of EM.

The second problem is that assurance against error through misidentification seems to depend (in part, at least) on our capacity to know certain features of our own judgments on the basis of ordinary self-knowledge. This suggests that the presence or absence of epistemic assurance against EM should be made to depend only on cases where the subject knows what normal self-knowledge allows us to know about our own judgments – cases of blatant self-ignorance should be screened out as irrelevant. Taking these two points into account, we arrive at:

(*Assurance*) Necessarily, for any judgment J, if J is of kind Φ, then, if the author of J knows what normal self-knowledge allows us to know about our own judgments, what he or she knows is incompatible with J being an error through misidentification.

Notice that *Assurance* is, just like *Impossibility*, a schematic principle: 'Φ' is a schematic letter that needs to be replaced by a kind term to obtain a true (or false) claim. In this way, we leave room for different hypothesis as to the nature of the kind in question (once again, the point will become relevant in the next section). Notice also that, unlike *Assurance*-1, *Assurance* does not require that, whenever the judgment is of kind Φ, the author of the judgment be in a position to know that his or her judgment is not an error through misidentification. It only requires that, whenever the judgment is of kind Φ *and the author of the judgment possesses adequate self-knowledge*, he or she know something which is *incompatible* with the judgment involving an EM.

(Two side points. I assume that there is such a thing as normal self-knowledge and that it makes sense to speak of "what normal self-knowledge allows us to know about our own judgments". In line with what I argued in the last chapter, I do not assume that there is a single way of knowing (be it introspection, or something else) that explains how we normally gain self-knowledge. But I do assume that there are things that normal self-knowledge allows

us to know and things that normal self-knowledge does *not* allow us to know. If I am thinking that p, I might be able to know, on the basis of normal self-knowledge, that I am thinking that p. I will never be able to know, on the basis of normal self-knowledge, what the neural correlate of my thought is (if there is such a thing).

Second point: there are various ways in which *Assurance* can be tweaked. For one thing, a somewhat different versions of the principle can be obtained by replacing "knows" with "is in a position to know" and "to know" with "to be in a position to know". For another thing, there are various ways to modify the notion of "what normal self-knowledge allows us to know". One can talk of what self-knowledge allows *normal subjects* to know in *normal circumstances*, but also also of what self-knowledge allows normal subjects to know in *optimal circumstances* or, more speculatively, of what self-knowledge would allow an *ideal subject* to know in *ideally optimal circumstances*. My preference goes to something along the lines of the second option: normal subjects in optimal circumstances. The other ways of construing the principle are equally possible and interesting, but a detailed discussion of their costs and benefits would take us too far afield).

Some might object that, as I formulated it, *Assurance* is too weak to solve the problem posed by ERROR-FREE and KNOWLEDGE. For example, consider again my judgment *that Fred is walking along the river*. Doesn't ordinary self-knowledge allow me to know, of my judgment, *that it is thus* (where 'thus' is some kind of type demonstrative that expresses the maximally specific way in which my judgment actually is)? It seems so. But being 'thus' is incompatible with being an EM (if 'thus' expresses the maximally specific way in which my judgment actually is, nothing could be 'thus' and be an EM). So, even if my judgment is *not* immune to error through misidentification, ordinary self-knowledge allows me to know something about it (namely, *that it is thus*) that is incompatible with the presence of any EM. This shows that immunity to error through misidentification cannot be reduced to the sum of

Impossibility with *Assurance.*

I think that one can respond to this objection by drawing a distinction between two propositions being mutually *inconsistent* and two propositions being mutually *incompatible*. Inconsistency can be defined as the contrary of compossibility: for p and q to be mutually inconsistent is just for p and q to exclude one another, i.e. to hold true in disjoint sets of possible worlds. Incompatibility, by contrast, can be understood as involving more than that: when p and q are mutually incompatible, anyone who knows p and is sufficiently trained with the concepts necessary to grasp q will be able to work out, from his or her knowledge of p, that q doesn't hold. Incompatibility is – one might say – manifest inconsistency, or something near enough. Applying this distinction to the case of my judgment *that Fred is walking along the river*, one could say the following: although what I know on the basis of self-knowledge is *inconsistent* with my judgment being an EM (for being 'thus' requires, among other things, not involving any EM), it is not *incompatible* with my judgment being an EM (because not even someone sufficiently trained with the concept of an EM could work out that my judgment is error-free from the mere observation that it is 'thus'). By contrast, when I judge *that I'm thirsty* on the basis of introspection, I seem to know something about my judgment that is incompatible, and not just inconsistent, with it being an EM – the absence of any error is fully (or almost fully) 'in view' for its author.

Admittedly, the notion of incompatibility at play here is somewhat vague and imprecise.[86] But I take it that it is clear enough for present purposes – mainly, to make *Assurance* less demanding than *Assurance*-1 without rendering it entirely vacuous. Different versions of *Assurance* can certainly be distinguished, depending on how liberal or strict we want to be about what propositions should count as mutually incompatible. The suggestion I

86 To make it more rigorous, we would have to be explicit about the resources one can draw on when one *works out* one piece of knowledge from another – can one appeal to some very general pieces of background empirical knowledge? Or is it only one's conceptual competence that one can rely upon?

want to put forward is simply that IEM arises when a kind satisfies both *Impossibility* and (some version of) *Assurance* – some hypothesis along these lines is needed to explain why ERROR-FREE and KNOWLEDGE do not constitute cases of IEM.[87]

1.3. What does it take to explain *Immunity*?

Now that we know (or at least have some credible hypotheses as to) what an error through misidentification is and what immunity to such errors consist in, we can get clearer on what it takes to *explain* the phenomenon Wittgenstein pointed out in the *Blue Book*, namely:

(*Immunity*) The judgments I form about my mental life are immune to error through misidentification.

Before we do that, however, one last objection needs to be addressed, concerning not so much my definitions of EM and IEM, but rather the way in which I proposed to formulate *Immunity* itself. The objection has to do with the following case. Suppose I learn via testimony that the oldest person in this room is angry and that, having independent reasons to believe that I am that person, I form the judgment *that I am angry*. Suppose further that I am wrong: although

87 The case of KNOWLEDGE raises another difficulty: philosophers who think that if one knows that p then one knows that one knows that p will say that KNOWLEDGE satisfies both *Impossibility* and *Assurance*. There are, I think, two ways of reacting to this. One possibility is to say that the philosophers in question are wrong – knowing that p does not entail knowing that one knows that p – and this explains why KNOWLEDGE is not a species of IEM. Another possibility is that, intuitions to the contrary notwithstanding, KNOWLEDGE *should* be regarded a species of IEM (or, at least, should be regarded as a species of IEM by philosophers who think that knowing that p is sufficient for knowing that one knows that p).

the oldest person in the room is, indeed, angry, I am not that person and I am not angry. In this case, one can plausibly say that my judgment *that I am angry* involves an error through misidentification. But that judgment is also a judgment about my mental life. So *Immunity* – as I proposed to formulate it – is not true, because at least some of the judgments I can form about my mental life are *not* immune to EM.

If we take this objection seriously, we might want to consider alternative ways of characterizing the phenomenon that Wittgenstein was interested in. For example, we might try the following:

(*Immunity**) The judgments I form about my mental life *purely on the basis of introspection* are immune to error through misidentification.

(*Immunity***) The judgments I form about my mental life *in a purely non-inferential way* are immune to error through misidentification.

But, while arguably these proposals succeed in ruling out the counterexample, they seem to invite confusion between two different tasks: the task of providing an agreed-upon *characterization* of the phenomenon Wittgenstein put his fingers on and the task of *explaining* that phenomenon. As we'll see in the next section, being based only on introspection and being formed in a purely non-inferential way are properties that (some) philosophers have appealed to in order to explain why certain judgments display IEM. Precisely because these properties have been taken to be part of the *explanans*, it would be nice if we did not have to mention them in the context of describing the *explanandum*.

This is possible if we respond to the objection in a different way: instead of trying to amend *Immunity*, one can simply suggest that *Immunity* should be understood as akin to a

generic claim. There is a difference between saying that *all* the judgments I form about my mental life are immune to error through misidentification and saying that *in general* they are. The latter reading of *Immunity* (which, incidentally, is much more consonant with the informal tone of Wittgenstein's remarks in the *Blue Book*) leaves room for uncharacteristic cases that violate the general rule, and the idea is that my judgment *that I am angry* (in the example described above) might be regarded it as one of them. Different philosophers will, of course, disagree on what features make that judgment an uncharacteristic case – some might point to the fact that it was not formed purely on the basis on introspection; others to the fact that it was not formed in a purely non-inferential way. But disagreement on this point will not disagreement on the truth of *Immunity* itself. So, by construing *Immunity* as generic, we can dismiss the alleged counterexamples as irrelevant, while remaining entirely neutral among competing explanations of the phenomenon we are trying to pin down.

Having explained why I do not think we need to replace *Immunity* with *Immunity** or *Immunity***, let me go back to my initial question in this section: what does it take to explain *Immunity*? Consider any claim of the following form:

The judgments I form about my mental life are immune to error through misidentification

because they belong to **Φ**

Given our discussion so far, there are at least three things we should require for a claim of this sort to count as a satisfactory explanation of *Immunity.* The first that it be true in general that the judgments I form about my mental life belong to **Φ** (insofar as *Immunity* is understood as a generic claim, its explanation, too, will have generic truth-conditions). The second is that **Φ** satisfy *Impossibility* and *Assurance* (if IEM involves not only the impossibility of any EM, but also the epistemic assurance of the author of the judgment against EM, **Φ** should satisfy

both principles at once). The third is that the claim be genuinely explanatory, in particular that it be informative and non-circular (if someone were to specify Φ as "whatever kind the judgments in question belong to that satisfies *Impossibility* and *Assurance*", the resulting claim would hardly constitute a satisfactory explanation of *Immunity*). In the rest of this chapter, I will distinguish different candidate explanations of *Immunity* and assess them in the light of these requirements.

2. *Immunity* and the Mainstream View of the Mental

In the last section, I clarified what an error through misidentification is, what immunity to such errors consist in and what it takes to explain the fact that the judgments I form about my mental life exhibit it. In this section, I will look at two prominent attempts to explain *Immunity* within the framework of MVM. Both these attempts accept the truth of *Particularity* and both of them explain *Immunity* in epistemological terms. Why these commonalities?

Particularity – it will be recalled – is the thesis that the judgments I form about my mental life are particular rather than general:

(*Particularity*) The judgments I form about my mental life are judgments about a particular individual, namely myself.

Now, it's not as though no defender of MVM ever denied *Particularity*. In fact, the thesis that judgments we form about our own mental life are mostly 'subjectless' judgments (i.e.

judgments that do not contain, or mention, or otherwise make reference to, any specific subject) has several illustrious supporters. Arguably, Hume subscribed to that thesis when he wrote that:

> For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other […]. I never catch *myself* at any time without a perception, and never can observe anything but the perception. (Hume 1978, 252)

But the thesis is more standardly associated with Lichtenberg, who said that:

> We know only the existence of our sensations, representations and thoughts. *It thinks*, we should say, just as one says, *it lightnings*. To say *cogito* is already too much if we translate it as *I think*. (Lichtenberg 2012, 152)

We know that, following the lead of Hume and Lichtenberg, Wittgenstein denied *Particularity* explicitly:

> Wittgenstein reportedly held at one time that "I have toothache" and "'He has toothache" are not values of a common propositional function, that in "I have toothache" the word 'I' does not "denote a possessor," and that "Just as no (physical) eye is involved in seeing, so no Ego is involved in thinking or in having toothache." He is also reported to have viewed with approval Lichtenberg's saying that instead of "I think" we ought to say "It thinks" (with "it' used as it is in "It is snowing"). (Shoemaker 1968, 555-556).

Some even went as far as denying *Particularity* to explain *Immunity*. On Elizabeth Anscombe's view, for instance, mental self-ascriptions (and other judgments we verbalize using the first-person pronoun 'I') are IEM precisely because, appearances to the contrary

notwithstanding, their content does not concern any object in particular: “getting hold of the wrong object *is* excluded [because] there is no getting hold of an object at all” (Anscombe 1975, 59).

Despite this venerable tradition, however, the combination of MVM with the denial of *Particularity* strikes many as a weird combination. The reason is easy to see. MVM does not discriminate mental states and MENTAL STATES. So, in the context of MVM, the only 'Lichtenbergian' propositions that the avowal “I am thirsty” could be taken to express are the propositions *that there is thirst* or *that it thirsts* (with “it” used as in “It is snowing”). But one noteworthy feature of these propositions is that they can be true when *I* am thirsty, when *you* are thirsty or when *anyone else* is thirsty – in other words, their truth is entirely independent of *who* is thirsty. In what sense, then, can one insist that judgments having 'Lichtenbergian' propositions as their content are judgments about what *my* mental life, as opposed to judgments about *anyone else's* mental life? There may well be intuitions supporting the denial of *Particularity.* But, against the backdrop of MVM, it is entirely unclear what to make of these intuitions.

Particularity, then, cannot be given up very easily, at least if MVM is true. And if *Particularity* is not given up, one is left with little choice but to explain *Immunity* in epistemological terms. For if the explanation of *Immunity* has nothing to do with *what* I judge – if the content of the judgments I form about my mental life resembles, in every relevant respect, the content of judgments that are routinely vulnerable to EM – it is natural to expect that it will have something to do with *how* I come to judge whatever I judge, i.e. with the particular epistemic resources I draw on to arrive at those judgments. The main options here are two: the explanation could appeal to the *grounds* on which the judgments in question are based or else to the *manner* in which they are formed. In this section, I will examine each of these two options in turn and explain why I find them unappealing.

The Simple Explanation

The first alternative, which is by far the most popular one, ties IEM with the *grounds* on which a judgment is based and is sometimes referred to as the "Simple Explanation" of *Immunity*:

> The *Simple Explanation* just consists in the observation that a judgment will be immune to error through *misidentification* when it is not based on an *identification*. The *Simple Explanation* may not appear very exciting- or perhaps even very explanatory. However, what it says does at least seem true. […] No one should deny the truth of the *Simple Explanation*. (Morgan 2012, 106)

Though hints of it can be found in Shoemaker (1968), the origins of the Simple Explanation are probably to be traced back to Gareth Evans's *The Varieties of Reference*. Evans begins by distinguishing two kinds of singular knowledge:

> 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*. […] We might say that knowledge of the truth of a singular proposition is *identification-free* if it is not identification-dependent. (1982, 180)

After showing that the distinction holds for judgments more in general, he claims that:

> Clearly, [identification-free judgments] are immune to a kind of error to which [identification-dependent] judgments are liable. *Since they do not rest on any identification, they are immune to error through misidentification*. (1982, 182; my emphasis)

In its broad outline, Evans's account has been accepted by many participants in the debate surrounding *Immunity*. Implicitly or explicitly, it has been defended or endorsed by Peacocke (1999, 270), Wright (1998; 2012), Coliva (2006), Nida- Rümelin (2011), Stanley (2011), Morgan (2012), and Garcia-Carpintero (forthcoming), among others. The reasons for this popularity are not difficult to see. The Simple Explanation is not only simple. As Wright observes, it also "liberates us from any need for metaphysical or semantic extravagance in trying to account for the phenomenon [of immunity to error through misidentification]. It is a pleasingly deflationary account" (Wright 2012, 255).

If we call "BASIC" the kind including all and only the judgments whose grounds do not involve any identification (judgments that are not "based" or do not "rest" on any identification or that cannot "be seen as a result of an identification" or that are "identification-free"), the Simple Explanation can be put as follows:

The judgments I form about my mental life are immune to error through misidentification because they belong to BASIC

Three questions are therefore crucial for the success of the Simple Explanation:

(i) Is it true that, in general, the judgments I form about my mental life belong to BASIC?

(ii) Does BASIC satisfy *Impossibility* and *Assurance*?

(iii) Is it genuinely explanatory to say that the judgments I form about my mental life are immune to error through misidentification because they belong to BASIC?

My focus here will be on (ii). One immediate problem is that not all cases of EM are cases where the evidence that justifies the subjects in making the judgment includes an

identification. Recall the example Pryor (1999) uses to establish this point:

> *The skunk* – I smell a skunky odor and see several animals rummaging around in my garden. Approaching closer and sniffing, I form the belief, of the smallest of these animals, *that it is a skunk in my garden*. The belief is mistaken. There are several skunks in my garden, but none of them is the small animal I see.

My judgment in *The Skunk* seems to be justified by identification-free evidence. Nevertheless, it is natural to classify it as a case of EM. This means that, if belonging to BASIC is purely a matter of being justified by identification-free evidence, BASIC does *not* satisfy *Impossibility* and the Simple Explanation is in trouble.

In principle, one could reply to this objection by denying that my judgment in *The Skunk* is justified by identification-free evidence: one could insist that there is a particular skunk such that my olfactory experience justifies me in believing, of the smallest of the animals rummaging around in my garden, that it is identical with *that skunk*. Alternatively, one could say that my judgment in *The Skunk* exemplifies a species of EM that the Simple Explanation is not interested in explaining – a deviant or derivative species of EM. But these replies are somewhat unconvincing. They ignore two basic lessons we learned in § 1. The first that the label "error through misidentification" should not mislead us into taking the presence of a false identification among the grounds of a judgment to be strictly speaking necessary for EM. The second is that, far from being 'deviant' or 'derivative', "the phenomenon of [identification-free misidentification] is epistemologically more central – and more interesting – than the phenomenon of *de re* misidentification", not least because "immunity to [this broader kind of misidentification] is a more basic and more rare epistemic status" (Pryor 1999, 286).

Friends of the Simple Explanation have not been completely insensitive to these

considerations. This is why most of them have responded to the objection in a different way. They have denied that belonging to BASIC is purely a matter of being justified by identification-free evidence. Instead, they have proposed to characterize BASIC in the following way:

a judgment belongs to BASIC if and only if it is a judgment whose grounds *and background presuppositions* do not involve any identification

The notion of *background presupposition* requires some unpacking. As we know, the grounds of a judgment are nothing else than the evidence that justifies the subject in making that judgment. By contrast, the background presuppositions of a judgment are supposed to be those tacit or implicit commitments that, if called into question, would make it rational to withdraw the judgement, because it is only against the backdrop of such commitments that one took oneself to have evidence for the judgement. To illustrate, the content of my visual experience is the evidence I have for judging *that there is table in front of me*, whereas the proposition *that my visual apparatus is working properly* is, arguably, a background presuppositions of my judgment (if someone were to tell me that my visual apparatus is not working properly I would no longer take the content of my visual experience to support my belief *that there is a table in front of me*).[88]

Why adopting this narrower characterization of BASIC? The crucial thought is that "[the background] presupposition of a judgment may in a particular case include an identification. Should that presupposition fail, the judgment at issue may suffer error through misidentification even if it is [supported by identification-free evidence]"(Wright 2012, 270).

88 The distinction between evidence and background presuppositions of a judgment is explicitly drawn by Coliva (2006), although (as she points out) hints of it can be found already in Pryor (2000) and Wright (2002).

My judgment in *The Skunk* is supposed to provide an illustration of this. According to Coliva (2006), if someone were to call into question the proposition *that this animal (I can now see) is the animal (in my garden) which is actually responsible for this odor I can smell*, I would be rationally obliged to withdraw my judgment. So that proposition – a proposition identifying the animal I am seeing with the animal whose odor I am smelling – should be regarded as a background presupposition of my judgment. The general hypothesis is that "error through misidentification […] will *always* depend on the presence of a false identification component as part of [either the grounds or the background presuppositions of] a given judgment" (Coliva 2006, 417).[89] Define BASIC so as to exclude any identification from *both* the grounds *and* the background presuppositions of the judgment and the Simple Explanation will be safe.

I think there is much to be said in favour of this way of developing the Simple Explanation. The distinction between grounds and background presuppositions of a judgment is central to much current theorizing in epistemology, and it is certainly a nice features of this proposal that it puts this distinction to new use, rather than responding to the objection in some *ad hoc* way. Nevertheless, there seem to me to be two problems with this strategy that friends of the Simple Explanation have either underestimated or entirely overlooked.

The first problem becomes apparent if we focus on the kind of reconstruction Coliva proposes of *The Skunk*. Consider again the proposition *that this animal (I can now see) is the animal (in my garden) which is actually responsible for this odor I can smell*. Saying that this proposition is a background presupposition of my judgment is tantamount to saying that, when I formed my judgment, I somehow presupposed (perhaps implicitly or tacitly) that there was a unique animal in my garden which was causally responsible for my olfactory experience (at any rate, that is what the definite description "*the* animal in my garden which is

89 The same move can be found in Garcia-Carpintero (forthcoming).

actually responsible for this odor I can smell" suggests, on one natural way of interpreting it). But why should I have presupposed such a thing? Could I not have presupposed (perhaps implicitly or tacitly) that my olfactory experience was collectively caused – or perhaps even overdetermined – by the odour of many distinct skunks? It certainly seems so. In fact, it seems that it would *not* have been very rational of me to presuppose that there was a unique skunk whose odor was responsible for my olfactory experience. And even if it was rational to presuppose such a thing, it would not have been very rational of me to let that presupposition play any remotely justificatory role in the formation of my judgment. But then why thinking that there actually was, among the background presuppositions of my judgment, any (true or false) proposition worth describing as an "identification"?

As far as I can see, the problem that these questions raise is rather general and does not have anything to do with the specific way in which Coliva formulated her proposal. The challenge that the Simple Explanation originally faced was that of accounting for identification-free cases of EM. Taking the identification to be part, not of the grounds of the judgment, but rather of its background presuppositions does not seem to be a satisfactory way of meeting that challenge. Friends of the Simple Explanation should – I think – come to terms with the idea that a judgment can be a case of EM even if its entire justificational architecture (and not just the part of it we associate with the judgment's grounds) is identification-free. It may be possible to do this without giving up the Simple Explanation. But, at the very least, BASIC has to be defined in even narrower terms than Coliva and Wright have suggested – how much narrower it is difficult to see.

This takes me to the second problem. It seems to me that, the narrower and more sophisticated our characterization of BASIC is, the greater chances BASIC will have to satisfy *Impossibility* without satisfying *Assurance*. This is already evident if we bracket for a moment the first problem and define BASIC in the way suggested above, i.e. as including all

and only judgments whose grounds and background presuppositions that do not involve any identification. The author of a judgment does *not* typically have access to what, from the objective point of view of epistemology, can be seen as the background presuppositions of his or her judgment. In particular, when one performs a judgment one may very well not be able to tell whether or not the background presuppositions of his or her judgments include an identification (at least not in general and not if one knows nothing of epistemology). So the risk looms large that some judgments may be BASIC (in the sense defined above) *without* their authors being epistemically assured against EM

The point can be seen most clearly in the case of my judgment *that I'm thirsty*. In § 1.3, I suggested that, while in making that judgment I cannot be sure to be right, something about my judgment tells me that I am not committing any error *through misidentification*: given what I know about my judgment on the basis of normal self-knowledge, there is simply no question of having mistaken someone else's thirst for mine. Now, finding out that this is so does not require much effort or concentration or reflection on my behalf: to convince myself of the fact, I need only cast a quick glance at the judgment I am performing. Certainly I do not have to evaluate elaborate hypotheses about the 'background presuppositions' of my judgment. If that's what I had to do in order to rule out the possibility of being committing an EM, I doubt that I would enjoy the kind of epistemic assurance against EM that I actually enjoy. The worry, then, is that, even if being BASIC makes for the impossibility of an error through misidentification, it does not make for any kind of epistemic assurance of the author of the judgment against EM: when one makes a BASIC judgment, what one knows may well be compatible with the judgment being an EM, even if one possesses adequate self-knowledge.

An advocate of the Simple Explanation can respond to this problem in at least three different ways.

The first would be to insist that the fact that a judgment is BASIC is, indeed, 'fully in

view' for its author, in the minimal sense that an ideal subject exercising self-knowledge in optimal circumstances would be in a position to know that fact. The observation would not give us a very strong kind of epistemic assurance. But perhaps it would give us *Assurance* of some kind. The question is whether we can content ourselves with such a weak version of the principle or whether we need a more robust sense in which we – imperfect subjects operating in less-than-optimal circumstances – are epistemically assured against EM when we make various judgments about what we believe, hope, desire, feel or experience.

The second response is more subtle and trades on the idea that, whenever a judgment is BASIC, its author may not know that this is so, but will know of some *other* feature of the judgment which happens to be incompatible with the judgment's involving an EM. Thus *Assurance* against EM is guaranteed, even for imperfect subjects operating in less-than-optimal circumstances. But the proponent of this line cannot get away with the vague suggestion that some feature or another of BASIC judgments allows their subjects to rest epistemically assured against EM. If it is agreed that it is not their being BASIC, one would like to know what the features in question are – in fact, one would expect the Simple Explanation to be re-formulated directly in terms of those features.

The third and last response would be to reject the considerations I offered in § 1.3 for regarding *Assurance* as part and parcel of IEM. This leaves the proponent of the Simple Explanation with only two options. The first is to show that, contrary to what I've argued there, IEM can and should be reduced to the mere impossibility of any error through misidentification. On this approach, in so far as it succeeds in showing that BASIC satisfies *Impossibility*, the Simple Explanation has succeeded in explaining all there is to explain. The second option is to offer some alternative hypothesis as to the additional ingredient that, together with *Impossibility*, makes certain judgments immune to error through misidentification.

This is not the place for a detailed examination of these responses. All I want to say is that – as things stand and given the two problems I outlined here – I find little justification for the optimism surrounding the Simple Explanation. I so agree that what the Simple Explanation says "does at least *seem* true", as Morgan puts it. But I also think that, when the notion of a judgment's being "based on" (or "resting on" or being "the result of") an identification is subjected to closer scrutiny, this appearance of truth vanishes and the prospects of the Simple Explanation start to look somewhat less promising.

The Very Simple Explanation

If it is not because of the *grounds* on which they are based, couldn't the judgments I form about my mental life display IEM because of the particular way in which they are formed? This is the approach advocated by Campbell (1999). He starts with the idea that one "could, in principle, have a way of finding out about particular properties which was, as it happens, confined to finding out about the properties of just one object" (1999, 93). Then he notes that "this way of finding out about properties might be fallible. But it would still be immune to error through misidentification" (ibid.). Finally, he suggests that this is what happens 'in the case of the first-person':

> In the case of the first-person, what is happening is […] that the subject is using ways of finding out about the world that are, as we might say, "dedicated" to the properties of one particular object, namely that very person. They are not ways of finding out that could be equally well applied to any range of objects. It is for that reason that although the subject using such a way of finding out can make a mistake, it could not be a mistake about who is in question. (Campbell 1999, 95)

All this should ring a bell. Campbell speaks of "ways of finding out about the world", but we could equally well speak of *ways of knowing*. He also speaks of a way of finding out about the world that is dedicated "to the properties of one particular object", but probably he has in mind a way of finding out about the world that is dedicated to the *mental* properties of one particular object (I presume that the way of knowing in question does not allow me to know things like my address or the name of my grandfather). If this is so, it is not too much of a stretch to read the passage above as speaking of what, in the last chapter, I called "introspection" – a way of knowing whereby one gains (direct) knowledge of one's mental properties. Campbell's main claim seems to be that, when a judgment is formed on the basis of introspection, this guarantees the impossibility of any misidentification, because the evidence one acquires through introspection either provides knowledge that one is so-and-so or it provides no knowledge at all. He doesn't offer any argument that, where there is introspection, there is also epistemic assurance of the author of the judgments against EM. But such an argument could readily be devised if it is assumed that when one forms a judgment on the basis of introspection, one is normally aware that one is doing so (i.e. one is normally aware that it is introspection that one is relying on).

The proposal is so simple that I will refer to it as the *Very Simple Explanation* of IEM. If we call INTROSPECTIVE the kind that includes all and only the judgments formed on the basis of introspection, the Very Simple Explanation could be put as follows:

The judgments I form about my mental life are immune to error through misidentification because they belong to INTROSPECTIVE

Although Campbell is, to my knowledge, the only philosopher who explicitly endorses a

version of the Very Simple Explanation of *Immunity*, many other philosophers have appealed to the Very Simple Explanation in passing or between the lines. So, for example, Recanati (2012b) suggests that what makes the judgments we form about our own mental states IEM is the particular way or 'mode' in which they are formed:

> Immunity to error through misidentification precisely arises from the fact that the subject which the judgment concerns is determined by the experiential mode in an invariant manner. Given the nature of the state, it can only concern the subject; there simply is no alternative, hence no possibility of error. (Recanati 2012b, 189)[90]

In effect, although I address it last, the Very Simple Explanation (or something along its lines) is probably the first thing that comes to mind when one is exposed to Wittgenstein's remarks about IEM: the reason why, in the case of the judgment one forms about one's mental life the possibility of a misidentification "has not been provided for" is that introspection only allows one to find out about one's own mental states, not about other people's mental states – could there be anything simpler than that?

Precisely because it is so 'catchy', it is important to see why the Very Simple Explanation is not a fully satisfactory explanation of *Immunity*. The success of the Very Simple Explanation depends on the answer to three distinct questions:

(i) Is it true that, in general, the judgments I form about my mental life belong to INTROSPECTIVE?

(ii) Does INTROSPECTIVE satisfy *Impossibility* and *Assurance*?

(iii) Is it genuinely explanatory to say that the judgments I form about my mental life

90 Here the Very Simple Explanation is given in terms of 'experiential' mode and elsewhere Recanati speaks of 'proprioceptive' mode, but these characterizations seem to be too restrictive. If it is to encompass the whole range of judgments about oneself that exhibit IEM, the best formulation of the Very Simple Explanation is the one outlined above, which is stated in terms of introspection

are immune to error through misidentification because they belong to INTROSPECTIVE?

The specific problems I will focus on here have to do mainly with (ii) and (iii), but let me note in passing that, given the 'pluralist' account of self-knowledge I defended in the last chapter, the answer to (i) is far from obvious.

As usual, let me start with *Impossibility*: the Very Simple Explanation says that that one cannot form a judgment on the basis of introspection and thereby commit an error through misidentification. Now I don't think anybody who accepts the existence of introspection should deny the truth of this observation. What should be questioned is the extent to which this observation is informative or illuminating. Notice that it is part of the meaning of the term "introspection" (as it is used here and in many other contexts) that either introspection provides knowledge about oneself or it provides no knowledge at all: "any process that generates knowledge equally of one's own and others' minds is by that token not an introspective process" (Schwitzgebel 2012a). A consequence of this convention is that, when a judgment about oneself fails to constitute knowledge and yet the minimal evidence on which it is based offers ∃-knowledge, that judgment cannot be called "introspective". Given our characterization of what an error through misidentification is (§ 1.2 above), it turns out we *defined* "introspection" as a way of knowing such that no judgment formed in that way is an error through misidentification – in brief, we defined "introspection" as an EM-proof way of knowing. This means that what the Very Simple Explanation says about *Impossibility* reduces to the following truism: one cannot form judgments about oneself through a certain EM-proof way of knowing and thereby commit an EM. True – as any truism – but not very helpful.

Of course, this is not to say that the Very Simple Explanation couldn't be developed into a more substantive account of *Impossibility*. In the context of a scientific investigation into the causes of planetary movements, one could define "gravity" as whatever force is

responsible for planets to move in ellipses. The claim that gravity causes planets to move in ellipses would then be a truism, but further research into the nature of gravity and further hypotheses about what it does could turn it into a more substantive thesis. The same might happen with introspection. For example, someone could follow Armstrong in identifying introspection with a certain type of "self-scanning process in the brain" (1968, 324). At that point, the central claim of the Very Simple Explanation with respect to *Impossibility* would be that one cannot form a judgment as a result of that type of self-scanning process and thereby commit an error through misidentification. Claims of this kind are open to all sorts of counterexamples and I am sceptical that they can be plausibly defended.[91] But at least they are substantive claims. If we stick to Campbell's original formulation, the Very Simple Explanation of *Impossibility* may be true, but it is entirely empty.

Even bigger worries arise in connection with *Assurance*. If we define introspection as a way of knowing that is "dedicated" to one's own mental properties, it (trivially) follows that no judgment formed on the basis of introspection involves any EM. But does it also follow that the author of any judgment formed on the basis of introspection is epistemically assured against EM? Clearly not. For, even if there is a way of knowing "dedicated" to one's own mental properties, no reason has yet been given why it should not be possible to form a judgment on the basis on that way of knowing *without* being aware that one is doing so (and, more in general, without being aware of any feature of one's judgment that is incompatible with the presence of EM). Hence, if we define introspection as a way of knowing that is "dedicated" to one's own mental properties, INTROSPECTIVE will satisfy *Impossibility*, but

91 The point is that the very same mechanism or process that (if Armstrong is right) is responsible for self-knowledge in the actual world might be responsible for other kinds of knowledge in other possible worlds. In particular, it's not difficult to imagine worlds where subjects are wired up in such a way that they (sometimes) acquire information about other minds using mechanisms and processes that, in the actual world, lead exclusively to self-knowledge. In this connection, see McGinn (2004).

it might well fail to satisfy *Assurance*. Even conceding that the answer to (i) is 'yes' (contrary to what a 'pluralist' view of self-knowledge would suggest) and even conceding that the answer to (iii) is 'yes' (despite the way "introspection" was defined), it is entirely unclear that the answer to (ii) can be 'yes'. So it is entirely unclear that the Very Simple Explanation can be made to work.

Let me sum up. I suggested that, if MVM is true, *Particularity* cannot be given up very easily and, if *Particularity* is true, *Immunity* can only receive some kind of epistemological explanation. Epistemological accounts of *Immunity* come in two varieties, depending on whether they appeal to the *grounds* (or *background assumptions*) on which the judgment is based or to the *way* in which it was formed. The most promising version of the first variety is the Simple Explanation. The most promising version of the second variety is the Very Simple Explanation. What I've been arguing is that, as things stand, neither the Simple Explanation nor the Very Simple Explanation offer a fully satisfactory account of *Immunity*. To be sure, this doesn't mean that if MVM is correct, the project of explaining *Immunity* is completely hopeless. Maybe the Simple and the Very Simple Explanation can be fixed. Or maybe other (less simple, perhaps, but more convincing) epistemological accounts of *Immunity* can be offered in their place. This is work for philosophers who are more invested in MVM than I am. In the next section, I will look at how *Immunity* can be explained if we give up MVM and embrace SVM instead.

3. *Immunity* and the Subjectivist View of the Mental

The kind of problems I raised in the last section for the project of explaining *Immunity* in epistemological terms are not specific to MVM: if subjectivists were to adopt the Simple or the Very Simple Explanation, they would face the very same difficulties.

The crucial point, however, is that subjectivists do not *have* to explain *Immunity* in epistemological terms. Friends of MVM are forced to look for an epistemological explanation of *Immunity* because of two combined reasons. The first is that if MVM is true, *Particularity* is true as well. And the second is that if *Particularity* is true, the content of the judgments I form about my mental life does not explain why they are immune to error through misidentification. One possibility remains open: MVM is *not* true, the judgments in question are *not* particular and so their content *does* explain why they enjoy IEM.

What kind of content would do for the purpose? In the last section I suggested that, if we take the judgments I form about my mental life to be 'Lichtenbergian', we cannot very comfortably take them to be about ordinary mental states – essentially because ordinary mental states do not discriminate between me and everyone else. Again, this seems to be the case whether or not MVM is true. The whole point is that if something along the lines of SVM is true there's an alternative to construing the judgments in question as judgments about ordinary mental states: they can be taken to be 'Lichtenbergian' judgments about MENTAL STATES.

So here is my proposal: *Immunity* is true because the judgments I form about my mental life are 'Lichtenbergian' judgments about MENTAL STATES. If we call 'LICHTENBERG' the kind including all and only 'Lichtenbergian' judgments about MENTAL STATES, the idea can be put as follows:

The judgments I form about my mental life are immune to error through misidentification because they belong to LICHTENBERG

The success of this account will depend on the answer to three questions:

(i) Is it true that, in general, the judgments I form about my mental life belong to LICHTENBERG?

(ii) Does LICHTENBERG satisfy *Impossibility* and *Assurance*?

(iii) Is it genuinely explanatory to say that the judgments I form about my mental life are immune to error through misidentification because they belong to LICHTENBERG?

In the remainder of this section, I will address each of these questions in turn.

(i) Is it true that the judgments I form about my mental life belong to LICHTENBERG?

Saying that the judgments I form about my mental life belong to LICHTENBERG is saying two things: that they are 'Lichtenbergian' judgments and that they are 'Lichtenbergian' judgments about MENTAL STATES. The first claim is one that (as we've already seen in the last section) many philosophers – from Hume and Lichtenberg to Wittgenstein and Anscombe – have been attracted by. The second claim echoes the hypothesis, defended in the last chapter, that self-knowledge is knowledge whereby one ascribes to oneself MENTAL STATES, rather than mental states. This connection will be further explored below. For the moment, the crucial thing to be noticed is that, by saying that the judgments I form about my mental life are 'Lichtenbergian judgments about MENTAL STATES, friends of SVM can retain a sense in which – even if they are general rather than particular – these judgments concern *me* as opposed to anyone else. For, according to SVM, I am the only subject who has MENTAL STATES.

So the circumstances in which I can correctly make a 'Lichtenbergian' judgment about a certain MENTAL STATE are all and only those in which *I* am in that MENTAL STATE. In other words, by bringing MENTAL STATES into play, friends of SVM can reject *Particularity*, while avoiding the counterintuitive consequences that rejecting *Particularity* has for someone who accepts MVM.

Other problems need to be addressed, however. In particular, the thesis that the judgments I form about my mental life are 'Lichtenbergian' judgments about MENTAL STATES can be objected to on semantic and psychological grounds. From a semantic point of view, I might be accused of flirting with the implausible theory that "'I' is neither a name nor another kind of expression whose logical role is to make a reference, *at all*" (Anscombe 1975, 60). From a psychological point of view, the obvious worry concerns the gap between the content of the judgments I form about my mental life – which is supposed to be general rather than particular – and the contents of full-blown self-knowledge – which concern me, as opposed to anyone else. What could possibly justify a transition from the general to the particular? How do we get full-blown self-knowledge from a 'Lichtenbergian' judgment?

These worries are legitimate, but not unanswerable. The first thing to be noticed is that one can follow Lichtenberg in taking the content of introspective judgments to be subjectless and reject Anscombe's controversial theory concerning the meaning of "I". On the hybrid view I have in mind, there is a mismatch between the general content of the judgments I form about my mental life and the particular content of the avowals used to verbalize them. Admittedly, this move creates a wrinkle in the account, but it's not as though no credible story could be offered of why we express our subjectless introspective thoughts in the way we do. For one thing, assuming the standard 'reflexive' account of the meaning of "I", the mere fact that a judgment *that it THIRSTS* is correct if and only if its author is thirsty explains why "I am thirsty" is not a completely inappropriate way of voicing the judgment. Strictly speaking, the

content of the judgment is *not* that *I* am thirsty, but, being the author of the judgment, I can make a correct assertion to the effect *that I am thirsty* whenever I can make a correct judgment to the effect *that it* THIRSTS. Moreover, the mismatch between the content of the judgment and the content of the avowal is something that any defender of SVM should *expect*, given that ordinary language doesn't contain a distinction between mental states and MENTAL STATES. It's not too much of a stretch to suppose that if that distinction were to find its way into ordinary language, we would feel more comfortable than we actually do to report how things are with our beliefs, hopes, desires, feelings and experiences in the way suggested by Lichtenberg.

The problem remains that there's a gap between the proposition *that it* THIRSTS and the proposition *that I am thirsty* (or, for that matter, the proposition *that I am* THIRSTY): that I am thirsty (or THIRSTY) is not a necessary consequence of the fact that it THIRSTS, let alone something that one could, as it were, 'logically deduce' from the fact that it THIRSTS. But it's not clear to me that this should be seen as a problem, rather than a virtue of the proposal under discussion. Suppose I am both thirsty and aware of my thirst. If the content of my awareness were *that I am thirsty* (or *that I am* THIRSTY) – or if the content of my awareness could be used to 'logically deduce' *that I am thirsty* (or *that I am* THIRSTY) – maybe there would be a way of coming to know or justifiedly believe that I exist on the sole basis of my awareness. But I agree with Hossack that there is no way one can come to know or justifiedly believe in one's existence on the sole basis of one's awareness of one's mental life:

> Mere consciousness cannot warrant my judgement that I exist. For although consciousness informs me of the existence of a state of what is in fact myself, it cannot inform me of the existence of my self [...]. Descartes missed this point in his famous paralogism *Cogito*, *ergo sum*. But his critics such as Hume and Kant did not miss it; they correctly insisted that mere consciousness of a conscious state does not on its own warrant a belief in the existence of

the subject of the conscious states. (Hossack 2006, 228)

Importantly, one can agree with Hossack without denying the obvious point that one's awareness of one's mental life can be used to come to know things about oneself. The fact that "mere consciousness" does not *on its own* warrant a belief in the existence of the subject of the conscious state does not mean that it doesn't do so in conjunction with other pieces of knowledge. Here is a sketch of how this might go if SVM is true. To get the account going, we need to make three assumptions. The first assumption is that it is a true and knowable fact that:

(1) From every point of view, only one subject is special (i.e. such that his or her mental states are MENTAL STATES)

The second assumption is that there is a certain distinctive kind of concept – the concept *I* – whose reference-fixing rule (or 'character', in Kaplan's (1989) terminology) is that:

(2) From every point of view, any THOUGHT involving the concept *I* refers to the subject who is special.

And the third assumption is that a subject who possesses the concept *I* has some implicit grasp of its reference-fixing rule (what 'implicit grasp' of a reference-fixing rule amounts to is, of course, a difficult question, but along with many other I assume we can make decent sense of this notion).

With these assumptions in place, let us suppose that, right now, awareness of my own mental life (or "mere consciousness", as Hossack calls it) tells me *that it THIRSTS.* I therefore

judge *that it* THIRSTS and know *that it* THIRSTS. Given my knowledge of (1), I can know that:

(3) From every point of view, it THIRSTS if and only if the subject who is special is THIRSTY.

And so, since it THIRSTS and I know that it THIRSTS, I can know that:

(4) The subject who is special is THIRSTY.

Now, notice that, if SVM is true, the reference-fixing rule associated with the concept *I* guarantees that, from any point of view, one THINKS correctly whenever one THINKS to oneself "I am the subject who is special, if anyone is". This in turn means that, if SVM is true, the reference-fixing rule associated with the concept *I* guarantees that no subject can think to oneself "I am the subject who is special, if anyone is" without one's thought being correct. Given that I have an implicit grasp of that reference-fixing rule, there is a clear sense in which, for me, the epistemic risk involved in thinking *that the subject who is special is* THIRSTY is neither higher nor lower than the epistemic risk involved in thinking *that I am* THIRSTY. To the extent I know *that the subject who is special is* THIRSTY, then, I can also know:

(5) I am THIRSTY.

Let me note a few things about this way of explaining the transition from 'Lichtenbergian' knowledge to self-knowledge.

First of all, it is important to see that the transition requires competence with a concept (i.e. the concept *I*) but also a piece of substantial metaphysical knowledge (i.e. knowledge of (1)). In my opinion, this is exactly what was to be expected. Ultimately, "mere consciousness"

provides me with knowledge about myself, but to expect that it will yield such knowledge immediately or with the sole help of 'logical deduction' is expecting too much. Recall Lichtenberg's warning: “We know only the existence of our sensations, representations and thoughts. *It thinks*, we should say, just as one says, *it lightnings*. To say *cogito* is already too much if we translate it as *I think*” (Lichtenberg 2012, 152).

Second, by explaining how one can gain self-knowledge (i.e. knowledge concerning oneself, rather than anyone in general) on the basis of 'Lichtenbergian' knowledge (i.e. knowledge that doesn't concern anyone in particular), the account connects nicely the thesis defended here – that the judgments I form about what I believe, hope, desire, feel, experience, etc, are 'Lichtenbergian' judgments concerning MENTAL STATES – to the hypothesis I defended in the last chapter – that self-knowledge is knowledge whereby one ascribes to oneself MENTAL STATES.

Third, if it is legitimate to assume that there is a distinctive kind of first-person concept with (2) as its reference-fixing rule, the account explains more than just how we can gain self-knowledge on the basis of 'Lichtenbergian' knowledge. It also has the potential to explain why most self-knowledge is knowledge whereby one ascribes to oneself MENTAL STATES *under a distinctive first-person mode of presentation*. Of course the claim that there is a distinctive kind of first-person concept with (2) as its reference-fixing rule requires more defence than I can give it here. But it is certainly an intriguing hypothesis, not least because of the striking analogy between (2) and the reference-fixing rule associated with the concept *Now*, which is generally taken to be that:

(2') At every time, any present thought involving the concept *Now* refers to the present instant.

Fourth and last, if it could be shown that not only there is a distinctive kind of first-

person concept with (2) as its reference-fixing rule, but (in some of its uses, at least) the word "I" expresses that concept, the connection between 'Lichtenbergian' judgments and the first-person avowals we use to verbalize them would be even tighter than was suggested earlier. Once again, I cannot here hope to defend or even just explore the thesis that (in some of its uses, at least) the word "I" expresses a first-person concept with (2) as its reference-fixing rule. I will limit myself to observing that one could defend this thesis without denying the orthodox view that "I" is a rigid designator. For, if (2) captures only the Kaplanian 'character' of the concept *I*, none of what I said implies that the concept *I* does not refer to myself in all possible worlds, i.e. that it is not *modally* rigid.[92]

(ii) Does LICHTENBERG satisfy Impossibility and Assurance?

The distinctive feature of a 'Lichtenbergian' judgment is that it is not a judgment about any particular object or entity. This being so, it is straightforward to see that LICHTENBERG satisfies *Impossibility*. The only real question is whether it also satisfies *Assurance*. Here is my argument for thinking that it does:

> (i) Necessarily, if one performs a judgment that belongs to LICHTENBERG and knows what normal self-knowledge allows us to know about our own judgments, one knows that one's judgment is not about any particular individual.
>
> (ii) If (i), then, necessarily, if one performs a judgment that belongs to LICHTENBERG and knows what normal self-knowledge allows us to know about

92 Incidentally, this marks a major point of difference with Hare's thesis that "the word "I" is synonymous with a modally and temporally nonrigid definite description that refers to whoever it is that has present experiences" (McDaniel 2012, 404).

our own judgments, what one knows is incompatible with one's judgment being an error through misidentification.

(iii) If (ii), then LICHTENBERG satisfies *Assurance*.

I take it that the only controversial premise of the argument is (i). Given how the notion of incompatibility was glossed in § 1.2, (ii) boils down to the claim that anyone sufficiently trained with the concept of an error through misidentification could work out, from the fact that the judgment is not about any particular individual, that it is error-free. This seems rather plausible. As to (iii), it is a trivial consequence of how *Assurance* has been characterized (see, again, § 1.2).

Let us focus on (i), then. Why think that someone who performs a 'Lichtenbergian' judgment about MENTAL STATES can know, on the basis of normal self-knowledge, that the judgment he or she is performing is not about any particular individual? My reason for thinking this is the following. First of all, I think that the judgments one forms about one's own mental life are 'Lichtenbergian'. Second, I think that Hume and Lichtenberg *knew*, of their 'Lichtenbergian' judgments, that they were 'Lichtenbergian' (or, more cautiously, that they were not judgments about particular individuals). Finally, I think that Hume and Lichtenberg knew this not on the basis of any elaborate philosophical thinking, but just on the basis of normal self-knowledge: they relied on nothing more than careful consideration of their own phenomenology to come to endorse the hypothesis (in my opinion, the *true* hypothesis) that the judgments one forms about one's own mental life are general, not particular. But this means that, when one forms a 'Lichtenbergian' judgment *and knows what normal self-knowledge allows us to know about our own judgments* (as Hume and Lichtenberg did), one knows that one's judgment is not about any particular individual. And this is pretty much all (i) says.

If this reasoning is correct, (i) is little more controversial than the very hypothesis that

the judgments one forms about one's mental life are general rather than particular. If it appears more controversial than that hypothesis, it's only because it can easily be confused with other, more dubious claims that have been made about the nature and extent of self-knowledge.

For one thing, (i) does *not* say that whenever we judge something, we are in a position to know what we are judging – a claim famously defended by McKinsey (1991, 2002).[93] (i) concerns what we can know on the basis of self-knowledge *when we do have self-knowledge.* It does not say anything about when, how or by whom self-knowledge can, in fact, be attained – no privileged access thesis is implicated.

For analogous reasons, (i) does *not* say that, whenever we make a 'Lichtenbergian' judgment, we are in a position to know that we are not making a judgment about any particular individual. Again, it is crucial to see that normal self-knowledge is neither universal nor equally distributed. If what I said above is right, normal self-knowledge allowed Hume and Lichtenberg to find out something about the judgments we form about our own mental life. Many other philosophers missed that point.

Finally, (i) does *not* say that whenever one judges something and knows what normal self-knowledge allows us to know about our own judgments, one knows whether or not one's judgment is about a particular individual. Arguably, this is not true (the judgment *that Santa Claus is coming to town* does not concern any particular individual, but it takes more than self-knowledge to know this). All (i) says is that whenever a judgment is a 'Lichtenbergian' judgment about mental states and its author does not lack self-knowledge, he or she knows that the judgment is general rather than particular. This claim may require more articulate and extensive defense than I can give it here, but, as far as I can see, there is no obvious reason to deny it.

93 Perhaps a good case can be made that if McKinsey's claim is true, (i) is true as well. I won't elaborate on this here.

(ii) Is it genuinely explanatory to say that the judgments I form about my mental life are immune to error through misidentification because they belong to LICHTENBERG?

One objection that is often levelled against the *explanatoriness* (rather than the adequacy) of the 'Lichtenbergian' account is that it does not generalize to other instances of IEM. For example, it is commonplace to hold that perceptual demonstrative judgments (e.g. judgment like "This is red", made on a perceptual basis, while looking at a red object) enjoy IEM (i.e. one can be wrong as to whether the object is red, but not about which object is red).[94] At the same time, no one takes the content of perceptual demonstrative judgments to be general rather than particular. And this means that the 'Lichtenbergian' account cannot explain the IEM of perceptual demonstrative judgments.

The objection is often taken to be a knock-down one (the following passages build into the 'Lichtenbergian' account the thesis that the word "I" does not refer, but never mind):

> The idea that Wittgenstein's 'as subject' uses of 'I' are somehow shown to be non-referential by their having IEM should have been 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. (Wright 1998, 20-21)

> [One of the arguments that the 'Lichtenbergian' account of *Immunity* is misguided is] usually that immunity to error through misidentification is common to both some uses of "I" and, for example, to perceptual demonstrative judgments, for which there is no temptation to think they may not refer. (Coliva 2006, 405)[95]

94 In fact, there are counterexamples to this claim (see, for instance, Campbell (1997, 69-70) and (1999, 96ss)), but let that pass.

95 See also Stanley (2011, 91).

What can be said in response to this objection? Two possibilities should be distinguished.

If Wright's and Coliva's main complaint against the 'Lichtenbergian' account is that it does not account for *all* cases of IEM, their complaint does not carry much force. In § 1.2, I argued that there is IEM where there is *some* kind that satisfies both *Impossibility* and *Assurance*. None of what I said there implies that there is just *one* kind satisfies both *Impossibility* and *Assurance*. Nor can it be assumed without argument that there is just one kind that satisfies *Impossibility* and *Assurance*. In fact, there is no good reason to think that there is just one kind that satisfies *Impossibility* and *Assurance* – given the hypothesis that IEM consists purely in the joint satisfaction of *Impossibility* and *Assurance*, IEM is best seen as a 'plural' phenomenon, having as many realizations as there are kinds that satisfy *Impossibility* and *Assurance*. So the observation that LICHTENBERG explains at most certain kinds of IEM and not others is hardly sufficient to prove the 'Lichtenbergian' account wrong. (Incidentally, the same can be said of the Simple and the Very Simple Explanation: the mere fact that neither BASIC nor INTROSPECTIVE can explain *all* kinds of IEM does not constitute evidence against these accounts).

But maybe Wright and Coliva have in mind a different kind of objection. Maybe their point is that, even if IEM is a 'plural' phenomenon, any account of the particular kind of IEM displayed by the judgments I form about my mental life should take it as a datum that that particular kind of IEM is "similar" (as Wright puts it) to the one exemplified by perceptual demonstrative judgments. And since the letter is clearly not content-based IEM, neither is the former.

If this is what the objection boils down to, my reply to it is very simple: I disagree that we should take it as a *datum* that the kind of IEM displayed by the judgments I form about my mental life is "similar" to the one exemplified by perceptual demonstrative judgments. I

think that the question of what other kinds of IEM are similar to or dissimilar from the particular kind of IEM relevant to *Immunity* is part of what is *at stake* in the controversy over *Immunity*: assuming that it is not content-based IEM (as the objectors urge us to do) would be tantamount to begging the question against the 'Lichtenbergian' account (just as assuming that it is content-based IEM would be tantamount to begging the question against the Simple or the Very Simple Explanation).

I conclude that the objection that the 'Lichtenbergian' account of *Immunity* does not generalize to perceptual demonstrative judgments is neither here nor there.

CHAPTER 5:

Experiential Knowledge

Experiential truths – truths concerning what it is like to undergo various kinds of experiences – are extraordinarily easy to know for most of us. While I sip my espresso, I know what it is like to taste coffee. As I look at the clear sky on a sunny day of January, I know what it is like to see something blue. Coming to know these things requires little effort or concentration or intelligence: most of us, most of the time know or are in a position to know plenty of things like these. But knowledge of many experiential truths is also foreclosed to a wide range of individuals – individuals who, for various reasons, might be described as *inexperienced*. Consider my case: I have never tasted Vegemite, so I don't know what it is like to taste Vegemite.[96] Or imagine a scientist who has spent all of her life in a black-and-white room, reading black-and-white books: that scientist wouldn't know what it is like to see something red.[97] Or take any human being whatsoever: not being bats, we human beings don't know what it is like to be a bat.[98] There is a tight connection between being inexperienced (in certain ways) and lacking experiential knowledge (of certain kinds). And the connection does not seem merely contingent. Rather, it seems that:

(*Inexperience*) For every kind of experience, there are certain conditions of inexperience C such that, necessarily, a subject who is in C does not know what it is like to have an

96 The example is Lewis's (1990).

97 See Jackson (1982).

98 See Nagel (1974).

experience of that kind.

Some conditions of inexperience are temporary: one day I'll taste Vegemite and thereby come to know what it is like to taste Vegemite. But others are perpetual: some congenitally blind persons will never know what it is like to see something red. In fact, certain conditions of inexperience are inescapable: plausibly I could not have been a bat instead of a human being, so it is literally impossible for me to know what it is like to be a bat. But whether it is contingent or necessary, temporary or perpetual, the striking feature of a condition of inexperience is that it necessarily prevents the acquisition of (certain kinds of) experiential knowledge. As long as one remains in a condition of inexperience, one cannot attain (certain kinds of) experiential knowledge:

> We cannot explain what red is to a blind man, nor can we make such things clear to others except by leading them into the presence of the thing and making them smell, see or taste the same thing we do, or, at the very least, by reminding them of some past perception that is similar. (Leibniz 1989, 24)

It seems natural to ask why this is so. Why is it that being inexperienced necessarily prevents one from acquiring experiential knowledge?[99] Experiential truths are nothing else than true propositions of a certain sort – true propositions concerning what it is like to undergo various kinds of experiences. But propositions can generally be communicated: why can't we communicate experiential truths to the inexperienced? Experiential knowledge is nothing else than knowledge of a certain kind – knowledge of experiential truths. But knowledge can generally be shared: why is inexperience such an insuperable barrier to the sharing of

99 In fact, some may doubt that this is a legitimate question. I will address such scepticism in the next section.

experiential knowledge? Learning about experience is just a particular case of learning and, in general, there are various ways of learning things (e.g. by perception, by deduction, by testimony, etc.): why no way of learning things allows the inexperienced to learn experiential truths while remaining inexperienced? These questions are all the more pressing if it is true that:

(*Objectivity*) Every experiential truth is an objective truth.

For being an objective truth is being a truth that holds from every point of view – in particular, an objective truth holds from the point of view of inexperienced subjects just as much as it holds from the point of view of experienced ones. So take any experiential truth p and take any inexperienced subject s: if p holds from the point of view of s, why can't s come to know it? If the truth is *there*, part of the complete and true story about how the world *objectively* is, why can't s learn it one way or another?

In this chapter, I will look at some attempts to answer these questions and raise some doubts about their chances of success (§ 2). Then I will sketch the contours of an alternative account of *Inexperience*, premised on the idea that, contrary to what *Objectivity* says, experiential truths are subjective, not objective (§ 3).

1. Inexperience

Inexperience says that, for every experiential truth, there are certain conditions of inexperience such that, necessarily, anyone who is in those conditions does not know the experiential truth in question.

It may be urged that this principle is trivially true because there are no such things as experiential truths. Experiential truths are supposed to be true propositions about what it is like to undergo this or that experience. Some will say that there is no good reasons to think that there are such propositions because knowledge of what it is like to undergo this or that experience is a philosophical chimera – this would be the position of an *eliminativist* about experiential knowledge.[100] Others will say that there is no good reasons to think that there are such propositions because experiential knowledge is not propositional knowledge, i.e. it is not knowledge of propositions – this would be the position of, e.g., an advocate of the so called *ability hypothesis* about experiential knowledge (i.e. the hypothesis according to which experiential knowledge consists only in the possession of certain skills or abilities).[101] In what follows, my working assumption will be that these positions are misguided: along with many others, I think that experiential knowledge exists and is propositional in nature. I have no argument to offer for the existence of experiential knowledge, but (given that the eliminativist option is not particularly intuitive) it's not clear to me that the burden of proof falls on those who affirm its existence and not on those who deny it.[102] Several arguments have been offered

---

100See, for instance, Churchland (1983).

101See Nemirow (1990) and Lewis (1990).

102Notice also that if experiential knowledge, while not itself propositional, is nonetheless *necessary* for some kind of propositional knowledge, the problem would remain of explaining:

> [*Inexperience**] For every truth p knowledge of which requires experiential knowledge, there are certain conditions of inexperience C such that, necessarily, a subject who is in C

that experiential knowledge (if there is such a thing) is propositional in nature,[103] but I will not go through these arguments here. Eliminativists about experiential knowledge and advocates of the ability hypothesis can read this chapter in a conditional spirit, as answering the question 'What should one say about the connection between being inexperienced and lacking experiential knowledge *if* experiential knowledge exists and is propositional in nature?'

(Caveat. Saying that experiential knowledge is propositional in nature is saying that every time we ascribe experiential knowledge what we are ascribing is, in fact, knowledge of some proposition or another:

(*Propositionality*) For x to have experiential knowledge is for there to be some experiential proposition that x knows.

It is important to see that accepting *Propositionality* is perfectly compatible with denying:

(*Uniqueness*) For every kind of experience, there is a unique proposition p such that for x to know what it is like to undergo an experience of that kind is for x to know p.

Since experiential knowledge is just knowledge, concerning a certain kind of experience, of what it is like to have that kind of experience, it is straightforward to see that *Uniqueness* implies *Propositionality*. But the converse does not hold. To see this, consider a particular

---

does not know p.

103Some of these arguments rest on specific hypotheses about the content of experiential knowledge. For example, if experiential knowledge is knowledge-how and knowledge-how is propositional in nature (as Stanley and Williamson (2001) have argued), it follows that experiential knowledge is propositional in nature. Other arguments are more general. For example it has been argued that experiential knowledge is the kind of knowledge that can be used in inferential reasoning and that eliminates epistemic possibilities – both of which are unequivocal marks of propositionality (see Loar (1997)).

kind of experience, like the experience of seeing red. *Uniqueness* requires the existence of a single proposition p_{RED} such that knowing what it is like to see red is knowing that proposition. *Propositionality*, by contrast, is not so demanding. For *Propositionality* does not rule the following option: although there is no single proposition such that knowing what it is like to see red consists in knowing *that* proposition, there is a range of different propositions such that knowing what it is like to see red consists in knowing *one or another* of those propositions. In other words, for all *Propositionality* says, it might be that every time we know what it is like to see red we do this by knowing a proposition, but there is no proposition such that every time we know what it is like to see red we do this by knowing that proposition. In this chapter, I assume the truth of *Propositionality*, but not that of *Uniqueness*. Indeed, many of the proposals I will look at – as well as the proposal that I will put forward myself in § 3 – are perfectly compatible with the falsity of *Uniqueness*).

So far, I've explained why I don't think that Inexperience is trivially true. Among those who agree with me that *Inexperience* is not trivially true, some might suggest that *Inexperience* is some kind of non-trivial analytic truth, i.e. a non-trivial truth that holds purely in virtue of the meaning of the constituent concepts (in much the same way as non-trivial mathematical truths hold true in virtue of the meaning of mathematical concepts, on at least some accounts of non-trivial mathematical truths). This suggestion strikes me as misguided. Maybe *Inexperience* would have been an analytic truth if experiential truths had been characterized as those truths knowledge of which cannot be attained by inexperienced subjects. But that is not the characterization of experiential knowledge I offered.[104]

104In fact, it is in tension with the characterization of experiential truths I offered, because not all truths whose knowledge is necessarily foreclosed to the inexperienced are experiential truths. Let p be a proposition that, as a matter of fact, no inexperienced individual actually knows (if there aren't any such individuals, then consider any true proposition p). The 'Fitchean' proposition *that p and no inexperienced individual knows that p* is not (or need not be) an experiential truth. But is is a truth knowledge of which is necessarily foreclosed to the inexperienced.

Experiential truths are truths concerning what it is like to undergo this or that kind of experience. That one cannot have knowledge of (some of) these truths if one is in a condition of inexperience (of a certain kind) does not seem to be an analytic truth. I concede that, in order to remove all suspicions of analyticity, something more has to be said about conditions of inexperience (in particular, "inexperience" had better not be defined as the conditions of someone who lacks experiential knowledge). I will take up this task below.

Even if *Inexperience* is not a trivially or analytically true principle, some will object that it is not a principle we should try to *explain*. There are two kinds of objectors I have in mind: the quietist and the brutalist. The quietist thinks that it is just a boring platitude that inexperienced subjects do not know experiential truths – nothing puzzling or mysterious about it. The brutalist, by contrast, thinks that there is a genuine puzzle here, but that we shouldn't expect to solve it: on her account, it is a brute fact that inexperienced subjects do not know experiential truths. I don't have much sympathy for brutalism and I hope it'll be conceded that, as philosophers, we should treat it as a position of last resort. As for quietism, it strikes me an unstable position. For either the quietist thinks that the boring platitude has a boring explanation or she thinks that it has no explanation. In the latter case, her position is really a form of brutalism in disguise. In the former case, her position presupposes a particular view about how *Inexperience* should be accounted for. But as will soon become clear, no view about how *Inexperience* should be accounted for is really boring or obvious or trivial.

Inexperience concerns experiential truths and experiential truths are truths about what it is like to undergo various kinds of experiences. The expression "what....is like" deserves some commentary. Generally, philosophers say that knowing "what it is like" to undergo a particular experience is knowing which *phenomenal character* or *quale* is possessed by that experience. But the notion of the experiential character or quale of an experience is hardly more clear than the notion of what it is like to undergo that experience – indeed, the two

notions can be seen as interdefinable (Williamson 1990, 48-48). I suspect there's not much one can do to clarify both notions at once – as Ned Block once wrote (paraphrasing Louis Armstrong) asking what qualia are is a bit like asking what jazz is: "If you got to ask, you ain't never gonna get to know" (Block 1978, 281).

One thing I should make clear in this connection is that I do *not* think we should expect an account of experiential knowledge to succeed or fail on the basis of what linguists will discover about the use of the expression "what...is like" in ordinary English. For example, some have proposed that a subject who is looking at something red can know what it is like to see red by knowing the proposition *that seeing red feels like this* (where "this" is supposed to demonstratively pick out the phenomenal character or quale of the experience one is currently having). While I am unconvinced by this account of knowledge of what it is like to see red (for reasons that will become clear in the next section), I do not think that its success is hostage to findings about the use of the expression "what it is like to see red" in ordinary English. This is because – and along with many others – I believe that "uses of the frame 'what...is like' in the context of the philosophy of consciousness are not "normal": i.e., they are not pieces of fully literal ordinary language, but rather involve some use of metaphor or idiom or jargon or code" (Hellie 2004, 336).[105] Since the publication of Nagel's seminal paper "What It Is Like To Be A Bat" (1974), these codified uses have become common currency in philosophy, but there are still significant differences in the way "what...is like" is used by different philosophers. On the use I will be making of it, "what...is like" is a very flexible and versatile frame. For example, on my understanding of this expression, knowledge of "what it is like to drink water" can also be described as "knowledge of what it is like to drink H_20" or "knowledge of what it is like to drink something that tastes exactly like water". More generally, knowing what F-ing is like *just is* knowing what Q-ing is like, for any Q and any F

105Notice that this is not Hellie's own view.

such that Q-ing and F-ing produce in us experiences with the same phenomenal character or quale.

Inexperience says that, for every truth concerning what it is like to undergo this or that experience, there are certain conditions of inexperience that are incompatible with knowing that truth. I did my best to clarify what I mean by knowing "what it is like" to undergo this or that experience. Let me now saying something about the notion of being *inexperienced*. What does it take to be inexperienced?

One answer that naturally comes to mind is that being inexperienced is a matter of not having undergone experiences of a certain kind. For example, one could say that having never tasted Vegemite is a condition of inexperience and it is because I am in this condition that I don't know what it is like to taste Vegemite. But there are well-known problems with this way of characterizing inexperience. For example, suppose someone were to create a molecule-by-molecule duplicate of me. At the moment of his creation, my duplicate would not have undergone any experience. But it seems intuitive to think that he would come into existence already equipped with a lot of experiential knowledge. For example, he would know, just as much as I do, what it is like to taste coffee. If this is so and *Inexperience* is true, being inexperienced can't be a matter of not having undergone experiences of a certain sort.

A more promising approach to defining inexperience is the following. I said earlier that not everyone thinks that experiential knowledge is propositional in nature: according to Lewis's (1990) and Nemirow's (1990) ability hypothesis, experiential knowledge is not knowledge of certain propositions, but possession of certain skills or abilities. For example, knowing what it is like to see red is not knowing this or that proposition specifying what it is like to see red, but possessing an ability to imagine or recognize red.[106] Now, even if the

106The right way of specifying the abilities in question is a matter of controversy. In the original formulation of the New Ability Hypothesis, "knowing what an experience is like is the same as knowing how to imagine having the experience" (Nemirow 1990, 495). Lewis (1983a) suggested that knowing what it is like might

'ability hypothesis' is false because experiential knowledge is propositional in nature, something in the vicinity of the 'ability hypothesis' might still be true, namely that being inexperienced is a matter of lacking certain skills or abilities – skills or abilities to imagine or recognize things. For instance, congenitally blind persons do not know what it is like to see red because they lack the ability to imagine or recognize red – lacking the ability to imagine or recognize red is a condition of inexperience. By contrast, having never tasted Vegemite is *not*, strictly speaking, a condition of inexperience, because a molecule-by-molecule duplicate of someone who knows what it is like to taste Vegemite would know what it is like to taste Vegemite, despite having never tasted Vegemite. What is true is that, in my case, having never tasted Vegemite has resulted in me lacking the ability to imagine or recognize the taste of Vegemite – a condition of inexperience that prevents me from knowing what it is like to taste Vegemite. The hypothesis, then, is that a condition of inexperience is the condition of someone lacking certain skills or abilities – skills or abilities to imagine or recognize certain things. If this hypothesis is correct, what *Inexperience* says, is that, for every kind of experience, there are certain skills or abilities to imagine or recognize things such that, necessarily, anyone lacking them does not know what it is like to have an experience of that kind.

One objection to this way of defining conditions of inexperience comes from Conee (1994). Conee argues that one can know what it is like to see red and yet lack the ability to imagine or recognize red. In a variation of Frank Jackson's (1982) famous thought-experiment, he asks us to consider the case of a scientist who, besides having spent all of her

---

coincide with possession of a *set* of abilities: "abilities to recognize, abilities to imagine, abilities to predict one's behavior by imaginative experiment" (131). Sometimes, it is suggested that knowing what it is like might require "abilities to remember" (Lewis 1990, 516). More recently, the New Ability Hypothesis has been reformulated thus: "Knowing what it is like to experience E is having the ability to distinguish imagining or having experience E from imagining or having any other experience" (Nenay 2009, 699). We don't need to enter into this debate here.

life in a black-and-white room, is also unable to visually imagine things:

> Mary has no visual imagination. […] [She] is released from her black and white confinement and sees something red for the first time. At that point, while she is intently gazing at the colour of red ripe tomatoes, it is clearly true that she knows what it is like to see something red. She has made an exciting discovery, 'Aha!', she might well exclaim. Yet she is unable to imagine anything. *A fortiori,* she is not able to imagine, remember, and recognize the experience [...] Hence, knowing what an experience is like does not imply having any such abilities. (Conee 1994, 139).

Conee's main point is that, if one is looking at something red, one may know what it is like to see red even if one lacks the ability to imagine or recognize red.[107] If this is correct and *Inexperience* is true, lacking the ability to imagine or recognize red can't be regarded as a condition of inexperience, contrary to what I suggested above.

In principle, one could reply to Conee's objection by insisting that, in so far as she lacks any ability to imagine or recognize red, Mary cannot properly be said to know what it is like to see red (Levin 2007). But I prefer a more accommodating line of response. According to this line of response, lacking the ability to imagine or recognize certain things is necessary, but not *sufficient* to count as inexperienced. To be inexperienced (in a certain way) one has to lack the ability to imagine or recognize things (of a certain sort) but also not be undergoing any experience (of a certain kind). Since Conee's Mary lacks the ability to imagine or recognize red things, but is actually looking at something red, she doesn't count as inexperienced. So we can say of her (compatibly with *Inexperience*) that she knows what it is like to see red. If we follow this line, what *Inexperience* says is ultimately this: for every kind of experience, there are certain skills or abilities to imagine or recognize things such that,

107Arguments to the same conclusion have been offered by Tye (2000) and Alter (2001).

necessarily, anyone lacking them and not having any experience of that kind does not know what it is like to have an experience of that kind. Unpacked in this way, *Inexperience* can hardly be made out to be an analytic truth. And insofar as it is true, that of trying to explain why this principle is true seems a perfectly legitimate philosophical project.

One last observation before we proceed. It is important not to confuse *Inexperience* with another principle that has received a lot of attention in recent years:

(*Unscrutability*) For every kind of experience, no truth concerning what it is like to undergo that kind of experience is a priori entailed by (or, in Chalmers' (2012) terminology 'scrutable' from) a proposition specifying the complete and true physical description of the world.

Inexperience does not entail *Unscrutability*. For example, you might hold that (contrary to what *Unscrutability* says) experiential truths *are* 'scrutable' from a complete and true physical description of the world, notwithstanding the fact that no inexperienced individual can know (all) of them. One might wonder how a defender of this combination of views would explain the fact that an inexperienced individual who knows the complete and true physical description of the world cannot deduce experiential truths from it. But notice that saying that the complete and true physical description of the world a priori entails any experiential truths is *not* saying that anyone who knows the complete and true physical description of the world could figure out experiential truths from the armchair. For example, it might be that experiential truths involve concepts that no inexperienced subject can possess – in which case an inexperienced subject will not even be able to entertain the relevant entailments, even if they are true and a priori knowable. (The crucial observation here is that a priori knowability entails nothing about how we acquire concepts – in particular, it is compatible with experiences, skills, abilities, etc. playing an essential 'enabling' role, as it is

sometimes said).

Symmetrically, *Unscrutability* does not entail *Inexperience*. For example, you might hold that (contrary to what *Inexperience* says) experiential knowledge *can* be attained by inexperienced subjects of any sort, notwithstanding the fact that there are no a priori connections between the complete and true physical description of the world and any experiential truth. The idea could be that experiential truths can be learned by inexperienced subjects in many ways (including, e.g., via testimony) but cannot be deduced a priori from physical truths, for example because, as a general rule, there are few (if any) a priori entailments between propositions belonging to different domains of discourse.

Someone who is interested in the project of explaining *Unscrutability* may well be a quietist or a brutalist about *Inexperience*. Symmetrically, someone who is interested in the project of explaining *Inexperience* may find unjustified the worries surrounding *Unscrutability* (for instance on the basis of a sweeping scepticism about the a priori). My concern in this chapter is solely with *Inexperience*. This is not because I find the *Unscrutability* uninteresting or not worth investigating, but because *Inexperience* strikes me as more basic or fundamental phenomenon (which, incidentally, may be the reason why some are tempted to be quietists or brutalists about it).

2. *Inexperience* and the Mainstream View of the Mental

Suppose MVM is correct. According to MVM the world is objectively the way it is and, therefore, experiential truths are, like all other truths, objective truths. If experiential truths are objective truths (i.e. if *Objectivity* is correct), how can we explain the fact that there are distinctive conditions of inexperience incompatible with knowing them? What kind of

epistemological principles governing experiential knowledge guarantee the truth of *Inexperience*?

Two very general hypotheses can be distinguished. According to the first hypothesis, inexperienced subjects can at least *entertain* the objective truths knowledge of which is experiential knowledge. Not so according to the second hypothesis.

Let us consider the first hypothesis first. If the truths knowledge of which is experiential knowledge are objective truths, they hold true from all points of view, including those of the inexperienced. And if, besides being true from the point of view of the inexperienced, they are also propositions that the inexperienced can entertain, the explanation of why the inexperienced do not know them must have to do with the fact that they are not related to those propositions in the way we are related to the propositions we know. For example, they might fail to be *justified* in believing those propositions (or, at least, they must lack the *right kind* of justification for believing them). The hypothesis could be that experiential knowledge requires some kind of *evidence* that, for some kind of reason, inexperienced subjects do not possess.

It might be suggested, for instance, that phenomenal knowledge requires *first-hand evidence*, evidence acquired through experience (or, perhaps, acquired through experience and stored in memory).[108] This is a natural way of interpreting the passage from Leibniz I quoted earlier:

> We cannot explain what red is to a blind man, nor can we make such things clear to others *except by leading them into the presence of the thing* and making them smell, see or taste the same thing we do, or, at the very least, *by reminding them* of some past perception that is similar. (Leibniz 1989, 24; my emphasis)

108For one way of glossing the distinction between first- and second-hand evidence see Lackey (2011).

Leibniz seems to suggest that evidence acquired via testimony (or via means other than experience and memory) does not work as evidence for experiential knowledge. The idea is both simple and intuitive: what makes experiential knowledge experiential is the fact that evidence acquired *without* the help of experience is, as one might put it, ineffective for attaining it. (Kant held a similar view about aesthetic knowledge – aesthetic judgments can only be grounded in aesthetic experiences of the subject: you cannot learn by testimony that something is beautiful).[109]

Simple and intuitive as it may sound, I am not very confident that this idea can be made to work. First of all, notice that it is rather difficult to give the idea a definite shape. To do this, one would have to pin down exactly what kind of evidence should be *allowed* for experiential knowledge and what kind of evidence should be *ruled out*. But, given what we said in the last section about *Inexperience*, this is no easy task. We need to rule out more than just evidence acquired via testimony (after all, we need to explain why the inexperienced can't come to know phenomenal truths via means other than testimony, e.g. by performing some kind of deductive reasoning). But we can't rule out all the evidence acquired via means other than experience and memory (for not every instance of experiential knowledge requires evidence acquired through experience and stored in memory – the experiential knowledge that my molecule-by-molecule duplicate has doesn't, for example). What we need to rule out exactly remains very much up in the air.

More importantly, even if we succeeded in giving this kind of account a definite shape, it seems to me that the account wouldn't be very *explanatory*. If we are interested in accounting for the truth of *Inexperience*, we cannot content ourselves with the observation that evidence acquired via testimony (or via means other than experience and memory) does

109"The approval of others affords no valid proof, available for the judging of beauty" (Kant 2007, 114).

not work as evidence for experiential knowledge – not even if the observation were correct. For the question still arises: *why* doesn't it work as evidence for experiential knowledge? Suppose someone suggested that evidence acquired in France does not work as evidence for mathematical knowledge. And suppose that, surprisingly enough, this suggestion turned out to be correct. It would still be natural to ask: why is that so? Similarly here: pending an answer to *why* evidence acquired via testimony (or via means other than experience and memory) does not work as evidence for experiential knowledge we are back to brutalism.

I take it that reflection on these difficulties will encourage the advocate of the Epistemological Strategy to pursue the second hypothesis I mentioned at the beginning: if conditions of inexperience are such an insuperable barrier to the acquisition of experiential knowledge by inexperienced subjects, the explanation must be that experiential truths are *beyond the reach* of any inexperienced intellect. Inexperienced subjects must somehow lack the resources to even *entertain* experiential truths. David Lewis expresses some sympathy for this idea:

> When we lack information, several alternative possibilities are open, when we get the information some of the alternatives are excluded. But a [...] peculiar thing about phenomenal information is that it resists this treatment. […] I can't even pose the question that phenomenal information is supposed to answer: is it this way or that? It seems that *the alternative possibilities must be unthinkable beforehand.* (Lewis 1990, 512; my emphasis)

The details of this account will depend heavily on how propositions are thought of. For, presumably, the explanation of why inexperienced subjects cannot entertain experiential truths will go through some feature or another of the propositions that those truths express.

In principle, one possibility is to model propositions rather coarse-grainedly, as *sets of possible worlds*. But if propositions are modelled as sets of possible worlds, will any of them

be a suitable candidate for the role of experiential truth? Earlier on, I mentioned the suggestion that learning what it is like to see red is learning *that seeing red feels like this* (where "this" demonstratively picks out the phenomenal character of one's current experience).[110] If propositions are modelled as sets of possible worlds, that proposition *that seeing red feels like this* is the same as the proposition *that seeing red feels the way normal subjects actually feel when they look at something red in normal conditions* (at least, if the subject employing the demonstrative is a normal subject looking at something red in normal conditions). Now, it seems highly plausible to think that inexperienced subjects *can* entertain the latter proposition. In fact, it seems highly plausible to think that inexperienced subjects can *know* that seeing red feels the way normal subjects actually feel when they look at something red in normal conditions. So our original supposition must have been wrong: learning *that seeing red feels like this* is *not* learning what it is like to see red. The worry is that this dialectic will generalize: for every candidate experiential truth p, we will have a hard time denying that the inexperienced can entertain by p, if p is modelled as a set of possible worlds. A view that individuates propositions in a rather coarse-grained manner, as sets of possible worlds, seems to sit at odds with the idea we are currently pursuing – the idea that some propositions about experience are beyond the reach of any inexperienced intellect.

That idea makes much more sense against the backdrop of a different account of propositions – the Russellian one. According to this account, propositions are complexes of objects and properties. The Russellian view can distinguish between the proposition *that seeing red feels like this* and the proposition that *that seeing red feels the way normal subjects actually feel when they look at something red in normal conditions*: the two propositions hold true in the exactly the same possible worlds, but only the first has the phenomenal character of a red experience (referred to by "this") as one of its constituents. Furthermore, it might be

110Perry (2006) puts forward a suggestion along these lines.

suggested that, in order to entertain a proposition in which a certain property P enters as a constituent, one has to be *acquainted* with P, where being acquainted with a property requires having the right sort of causal interaction with it or being capable of discriminating it from other properties or knowing that it exists or satisfying some other combination of causal and/or epistemic conditions involving it.[111] Building on this suggestion, someone who adopts the Russellian account could put forward the following hypothesis: the propositions knowledge of which is experiential knowledge are propositions having various phenomenal properties as constituents and the inexperienced cannot entertain them because they are not acquainted with those properties.[112]

The proposal is no doubt intriguing, not least because, when Russell first introduced the notion of acquaintance, he used it precisely to characterize our unmediated knowledge of sense data – and what are phenomenal characters if not the contemporary successors of sense data? Moreover, one might think that there is something intuitive in the idea that inexperienced subjects cannot entertain propositions having (certain) phenomenal characters among their constituents – using a metaphor one could say that the thoughts of someone who doesn't know what it is like to see red cannot 'latch onto' phenomenal redness in the appropriate way.

For all its intuitiveness, however, the proposal seems to me to suffer from two major difficulties.

The first is that it seems to overgeneralize. For example, the proposition *that phenomenal redness is phenomenal redness* has phenomenal redness as a constituent. If the Russellian account of *Inexperience* were correct, being inexperienced would be incompatible

111More about this shortly. The idea that knowing a proposition having certain objects or properties among its constituents requires acquaintance with those objects and properties has its critics. See, for instance, Hawthorne and Manley (2012).

112See Conee (1994).

with knowing that proposition. But this does not seem to be the case: one does not need to have the ability to imagine or recognize red, nor to actually have an experience as of red to know *that phenomenal redness is phenomenal redness*.

Secondly, there are problems in specifying what acquaintance with a property requires. Here we are confronted with a dilemma: either we follow Russell in setting the bar for acquaintance *really* high or we follow most contemporary theorists of acquaintance in setting it *reasonably* high. If we follow Russell, we end up with a view on which we are acquainted with only a few properties and objects: (some) phenomenal properties, ourselves and a bunch of universals. This view strikes many as implausible, for it forces us to recognize a difference where, intuitively, there is none. For example, there does not seem to be any deep semantic, psychological or metaphysical divide between the proposition *that phenomenal redness is phenomenal redness* and the proposition *that being an electron is being an electron*. But there would be if we were acquainted with phenomenal redness and not with the property of being an electron, as Russell held.

On the other horn of the dilemma, suppose we relax the standards for acquaintance, while still being reasonably demanding: acquaintance with P requires knowing a great deal of things about P and having a great deal of causal interaction with instances of P. Then we are left with no account of *Inexperience*. For, no matter how inexperience is defined, there is virtually no limit to how much causal interaction inexperienced subjects can have with instances of phenomenal redness (we can even imagine that they be endowed with ultra-reliable devices for detecting phenomenal-redness instances in other subjects) nor is there any limit to how much they can know about it (we can even imagine that they know, for each and every property like phenomenal redness, the unique Ramsey-sentence that that property satisfies according to the complete and true theory of reality). So if the set the bar for acquaintance with phenomenal properties only *reasonably* high, we are no longer entitled to

suppose that all inexperienced subjects lack acquaintance with various phenomenal properties. Which means that we are left with no explanation of why *Inexperience* holds.

(Perhaps a third possibility should be considered: the bar could be set really high for acquaintance with phenomenal properties, but only reasonably high for all the other properties. Wouldn't this give us an explanation of *Inexperience* without committing us to the implausible consequences of Russell's original views about acquaintance? Not really. The question would rather receive a new formulation: why are the criteria for acquaintance with phenomenal properties so much more demanding than the criteria for acquaintance with all the other properties? If it is not meant to be merely a redescription of the problem, the hypothesis that this so because different varieties of acquaintance are at stake sounds entirely *ad hoc*).

So, to sum up, within the framework of a Russellian theory of propositions it is hard to see why the inexperienced should be incapable of entertaining phenomenal truths: one can appeal to the hypothesis that (different kinds of) inexperienced subjects are not acquainted with (different kinds of) phenomenal properties, but, given a reasonable conception of acquaintance, this hypothesis appears rather implausible (not to mention the fact that, if it were true, it would make more propositions unknowable to the inexperienced than it is intuitively desirable). This simple line of reasoning puts pressure on the proponents of the Epistemological Strategy to abandon the Russellian conception and start thinking of propositions in a broadly Fregean way.

On the Fregean conception, propositions are not bare complexes of objects and properties, but encode also specific modes of presentation under which the relevant objects and properties are represented.[113] To illustrate, the Fregean conception distinguishes the

113In fact, at least two distinct kinds of Fregeanism can be distinguished. According to the first, propositions encode not only a given combination of objects and properties, but also a particular conceptual guise or manner of representing that combination. According to the second, proposition are bare Russellian complexes

proposition *that Hesperus is bright* from the proposition *that Phosphorus is bright* on the ground that the first represents the planet Venus as Hesperus, whereas the second represents the planet Venus as Phosphorus. In order to know (or even just entertain) a proposition involving a certain mode of presentation one has to possess the corresponding concept (for example, one needs the concept *Hesperus* to represent the planet Venus as Hesperus and the concept *Phosphorus* to represent the planet Venus as Phosphorus). So someone who possesses the concept *Hesperus* but not the concept *Phosphorus* could know the proposition *that Hesperus is bright* without knowing the proposition *that Phosphorus is bright*.

Now, suppose we apply this approach to experiential truths, i.e. suppose we treat experiential truths as true Fregean propositions. And suppose we assume, with Locke, that there are certain kinds of concepts that one cannot possess if one finds oneself in a condition of inexperience (I take the 'ideas' Locke talks about in this passage to be mental representations akin to concepts):

> I think it will be granted easily that if a child were kept in a place where he never saw any other but black and white till he were a man, he would have no more *ideas* of scarlet or green, than he that from his childhood never tasted an oyster, or a pine-apple, has of those particular relishes. (Locke 1975, 107)

Then the following hypothesis will naturally suggests itself: inexperienced subjects lack experiential knowledge because they lack the concepts one needs in order to know (or even just entertain) experiential truths. For example, a child who spent all of his life in a black-and-white room ("a place where he never say any other but black and white") would lack the

---

of objects and properties and various propositional attitudes (hoping, believing, desiring, knowing, etc.) encode a three-place relation whose relata are the subject of the state, the proposition and a guise or manner of representing it. The difference between these two forms of Fregeanism will be overlooked in what follows.

concepts one needs in order to know what it is like to see something scarlet or something green. So he wouldn't know what it is like to see something scarlet or something green.

The hypothesis is not without intuitive appeal, but calls for some fine-tuning. For one thing, we don't need to follow Locke in thinking that what the inexperienced lack are color concepts ("ideas of scarlet or green") or other kinds of *sensory concepts* (i.e. concepts expressing sensory properties) – the hypothesis is best formulated in terms of *phenomenal* concepts, i.e. concepts expressing the phenomenal properties associated with various kinds of sensory and non-sensory experiences. Secondly, the proposal need not be that the inexperienced can't *possess* the relevant concepts – it could also be that they *either* do not possess those concepts *or* cannot exercise them (at least, not 'fruitfully', i.e. not in ways that result in the acquisition of experiential knowledge – more about this in a moment). Thirdly – and most importantly, perhaps – the proposal need not be that the inexperienced cannot possess or fruitfully exercise *any* of the relevant concepts. What might be said is that they cannot possess or fruitfully exercise *all* of those concepts (formulated in this way, the Fregean approach has a clear advantage over the Russellian one, for it allows us to concede that there are at least *some* conceptual guises under which the inexperienced can 'latch onto' phenomenal properties). So, in its most general form, the Fregean hypothesis about *Inexperience* could be this: inexperienced subjects lack experiential knowledge because experiential knowledge requires specific kinds of phenomenal concepts that inexperienced subjects either cannot possess or cannot (fruitfully) exercise.

At this point, the question is: why can't the inexperienced possess or (fruitfully) exercise the phenomenal concepts in question? In principle, a defender of the Fregean hypothesis could try devise a univocal answer to this question. But – given how many varieties of phenomenal concepts have been identified and distinguished in recent years – a 'divide-and-conquer' strategy seems most advisable. One can suppose that, for every kind of

experience, knowing what it is like to undergo an experience of that kind is a matter of knowing one or another out of a range of Fregean propositions differing from one another for the modes of presentation they involve.[114] More precisely, one can say that:

- Some experiential truths require *recognitional* phenomenal concepts in order to be known (or even just entertained). Recognitional phenomenal concepts are phenomenal concepts whose possession involves or is constituted by a special ability to work out their extension "from an introspective perspective" (Levin 2007, 89) or "in perceptually (as opposed to intellectually) presented scenarios" (Yablo 2002, 461) or "in our phenomenological reflections" (Loar 1997, 601). As we've seen in § 1, being inexperienced (in a certain way) is partly a matter of lacking the skill or ability to imagine or recognize things (of a certain kind). Now, it might be suggested that the skills or abilities that the inexperienced lack are precisely those that recognitional phenomenal concepts require (or, at least, are necessary for those that recognitional phenomenal concepts require). Which is just to say that if one is inexperienced (in a certain way) one will lack recognitional phenomenal concepts (of a certain kind) and, consequently, knowledge of the corresponding experiential truths.

- Some experiential truths require *experience-involving* phenomenal concepts in order to be known (or even just entertained). Experience-involving phenomenal concepts are phenomenal concepts whose tokening involves having an experience of the kind the concept refers to (this could be put in more suggestive terms by saying that every token of the concepts 'quotes' the phenomenal property it refers to).[115] Now, as we've seen in § 1, being inexperienced (in a

114That is, one can suppose *Propositionality* to be true without *Uniqueness* being true (see § 1 above).

115Block (2007) has suggested something along these lines: "We could take the form of a phenomenal cognitive mode of presentation to be 'the experience:_' where the blank is filled by a phenomenal property" (264). The suggestion echoes some ideas explored in Chalmers (2003): "The referent [i.e. the experience] is somehow

certain way) is partly a matter of not undergoing any experience (of a certain kind). This means that if one is inexperienced (in a certain way) one can neither exercise experience-involving concepts (of a certain kind) nor know the corresponding experiential truths.

- Some experiential truths require *demonstrative* phenomenal concepts in order to be known (or even just entertained). If, in general, a demonstrative concept picks out whatever object or property stands in a certain relation to the subject (Recanati 2006; 2012a), a demonstrative phenomenal concept picks out whatever phenomenal property (or properties) one is currently instantiating.[116] Again, as we've seen in § 1, being inexperienced (in a certain way) is partly a matter of not undergoing any experience (of a certain kind). So although someone who is inexperienced (in a certain way) may be able to deploy phenomenal demonstratives, he will not be able to deploy them fruitfully, i.e. in ways that result in the acquisition of experiential knowledge (of a certain kind): his phenomenal demonstratives will 'misfire', so to speak. Which is just to say that if one is inexperienced (in a certain way) one will not know (certain kinds of) experiential truths involving demonstrative modes of presentation of phenomenal properties.

If it is assumed that every experiential truths requires recognitional or experience-involving or demonstrative phenomenal concepts, the Fregean account just canvassed will provide a neat explanation of *Inexperience*. In fact, a good case can be made that the 'Fregean' approach represents the most promising way in which an advocate of MVM can deal with *Inexperience*.

First of all, it seems prima facie natural to appeal to conceptual guises and modes of

present inside the concept's sense, in a way much stronger than in the usual cases of direct reference" (233).

116Notice that there is a complication here arising from the fact that, at any time, one is instantiating many determinable phenomenal properties, so it is unclear which of these properties a phenomenal demonstrative will pick out.

presentation to explain why *Inexperience* holds. After all, *Inexperience* is concerned with the role that abilities and experience play in making a certain kind of epistemic progress possible and one of the central theoretical functions of conceptual guises and modes of presentation is precisely to model and account for various kinds of epistemic progress (e.g. the epistemic progress made by someone who already knows *that Hesperus is Hesperus* upon learning *that Hesperus is Phosphorus*).

Second, saying that the inexperienced lack experiential knowledge because they cannot entertain experiential truths and that they cannot entertain experiential truths because they lack or cannot exercise certain kinds of experiential concepts seems a better explanation of *Inexperience* than just appealing to the fact that experiential knowledge requires evidence that the inexperienced do not possess: we are familiar with the idea that concept-possession and concept-exercise come with certain requirements, whereas the idea that, for certain epistemic purposes, evidence acquired in certain ways may be completely disqualified sounds somewhat unnatural and raises more problems than it solves.

Third, the 'Fregean' approach is more versatile than any acquaintance-based approach and, as already noted, it allows us to avoid some of the implausible consequences that a Russellian construal of acquaintance commits us to (in particular, it does not commit us to deny that inexperienced subjects can 'latch onto' phenomenal properties).

Fourth, and perhaps most important, there are good independent reasons to allow for the existence of many distinct modes of presentation under which one and the same phenomenal property can be conceptualized.

One reason is that we have conflicting intuitions about knowledge ascriptions involving inexperienced subjects. For example, consider again Frank Jackson's (1982) Mary, a neuroscientist who has read a great deal about phenomenal redness, but has never seen anything red and lacks the ability to imagine red, because she has spent all of her life in a

black-and-white room. There are intuitive grounds for thinking that Mary knows that it's phenomenal redness that normal human beings instantiate when they look at red objects in normal circumstances (she would answer 'Normal human beings instantiate phenomenal redness when looking at red objects in normal circumstances' when asked the appropriate question, her intention in saying so would be to express the belief that is normally expressed using those words and the language community she belongs to is – we may suppose – one in which the above sentence is used to claim that normal human beings instantiate phenomenal redness when looking at red objects in normal circumstances). On the other hand, there's also a strong intuition that someone like Mary does *not* know that it's phenomenal redness that normal human beings instantiate when looking at red objects in normal circumstances, because she doesn't know what it is like to see red. Both intuitions can be accommodated if one distinguishes various modes of presentation under which phenomenal redness can be conceptualized, for then one can say that Mary masters some of those modes of presentation but not others.[117]

Another good reason to buy into the 'Fregean' account sketched above has to do with the principle of *Unscrutability*, that we encountered earlier:

(*Unscrutability*) For every kind of experience, no truth concerning what it is like to undergo that kind of experience is a priori entailed by (or, in Chalmers' (2012) terminology 'scrutable' from) a proposition specifying the complete and true physical description of the world.

Some have argued that *Unscrutability* is in tension with *Physicalism*, the thesis that everything (including experiential facts) is or reduces to the physical.[118] But others have

117I report here, almost verbatim, the line of reasoning defended by Nida- Rümelin (1998, 53-55).

118Most notably, Chalmers (1996)

replied that the tension between *Unscrutability* and *Physicalism* is only apparent and that, once we recognize the existence of various kinds of phenomenal concepts, we can make our peace with the idea that truths involving those concepts are necessitated by the complete and true physical description of the world, even if they are not a priori entailed by it.[119]

Overall, it's undeniable that the 'Fregean' approach to *Inexperience* has many virtues. Yet there remain a cluster of worries about it. I will divide them into three groups: *theoretical* worries, *trivialization* worries and *intuitive* worries.

Theoretical worries – I said that modelling and accounting for various kinds of epistemic progresses is one of the chief theoretical functions of conceptual guises and modes of presentation. It is by no means their only function, however. In general, concepts are also used to make rational sense of the behaviour of a subject, to describe the economy of his or her thinking and, even more important, to account for the possibility of communication and agreement or disagreement between different subjects: concepts are the 'common currency' in which we trade thoughts with one another. It is partly because of the central role they play in communication that concepts can be possessed *deferentially*: it is always possible for someone to possess a concept by having the right kind of social interactions with a linguistic community whose practices (partially) fix the concept's extension and, in particular, by deferring to the experts of that community (this is what we learned on the basis of a number of examples made famous by Hilary Putnam (1970) and Tyler Burge (1979, 1986)). Now, some of the concepts that feature in the 'Fregean' account of *Inexperience* cannot be possessed deferentially: for example, one cannot possess a recognitional phenomenal concept deferentially, because recognitional concepts are concepts that one cannot possess unless one has certain abilities to imagine things. In principle there is nothing wrong with this:

119See, for instance, Loar (1997) and Block (2007).

phenomenal concepts could be the exception that proves the rule. And yet, as Ball (2009) argues at some length, this move has considerable theoretical costs. For example, we cannot use recognitional concepts to describe the content of most of the beliefs that Mary forms when she leaves the room, nor the kind of disagreements she can have with thinkers outside the black-and-white room or the kind of thoughts that thinkers outside the room form when they speak to Mary.[120] The general point is that, not being 'common currency', some of the concepts that feature in the 'Fregean' account can't be used to do all the things that concepts are normally used to do. Unlike Ball, I don't take this to be conclusive evidence against the existence of the concepts in question. But I do take this 'anomaly' to suggest that there's something *ad hoc* in the way concepts have to be used in the 'Fregean' approach to *Inexperience* and that, pending conclusive evidence that there are such things as recognitional concepts, it would be nice if we could find an alternative account.

(Some advocates of the 'Fregean' approach might be tempted to respond to the foregoing dialectic by conceding that inexperienced subjects can possess the relevant phenomenal concepts *partially* or *deferentially*, while insisting that they cannot possess them *fully* or *non-deferentially* and, therefore, lack full or non-deferential knowledge of the propositions in which those concepts feature.[121] This reply strikes me as unconvincing, for a very simple reason. The intuitive and, I would say, nonnegotiable starting point of any discussion about *Inexperience* is a point concerning *knowledge*: there are experiential truths

120For example, if 'red*' expresses a concept that Mary acquires upon seeing red for the first time, she cannot look back and truthfully think to herself

[1] I used to wonder what it's like to see red*, but now I know

For, as Ball observes, "it is hard to see how [1] could be true if what Mary knows involves a concept that she did not possess in her room, since the complement clause ellipsis in [1] eliminates the possibility that 'what it's like to see red' is ambiguous in the two attributions" (2009, 952).

121Ball (2009) discusses a reply along these lines and ascribes it to Chalmers.

that certain individuals are systematically prevented from knowing. Conceding that inexperienced subjects can know the relevant truths – that they can *know* what it is like to see red or taste coffee, albeit only in a 'partial' or 'deferential' fashion – is *not* a way of accommodating that point: it is a way of giving it up. There's nothing to be said about a view on which Mary *knows* what it is like to see red, except that it is a *misguided* view).

Trivialisation worries – A more serious worry about the 'Fregean' approach is that it runs the risk of *trivializing* the epistemic progress one makes when one acquires experiential knowledge. Taking Frank Jackson's Mary as a study case, we want to say that Mary comes to know (or is in a position to come to know) a *rather specific* fact when she sees red for the first time, a fact concerning what it is like to see red. And we also want to say that the kind of epistemic progress Mary makes when she sees red for the first time is a rather *extraordinary* epistemic progress, the kind of epistemic progress that one doesn't make every day. The problem is that it's not clear that the 'Fregean' approach allows us to say these two things.

The 'Fregean' approach predicts that Mary will be in a position to know an awful lot of things when she acquires the concepts she lacks. For instance, if *phenomenal redness** is a concept of phenomenal redness that Mary lacks when she's inside the room, it is only upon leaving the room that she'll find out *that phenomenal redness* is phenomenal redness** (in fact, the account says that it is only upon leaving the room that Mary will start to be able to *entertain* the propositions *that phenomenal redness* is phenomenal redness**). But it seems wrong to suggest that the sense in which Mary doesn't know (or cannot make hypotheses as to) what it is like to see red is also one in which Mary doesn't know (or cannot make hypotheses as to) whether the proposition *that phenomenal redness* is phenomenal redness** is true. The two cases are intuitively different, but the 'Fregean' approach is bound to treat them on a par – a problem faced also (albeit in a slightly different form) by the 'Russellian'

account explored earlier.[122]

The risk of trivialization is even more serious if, rather than being artificially restricted to the treatment of concepts of phenomenal properties, the 'Fregean' approach is extended also to other kinds of concepts. For example, suppose we say that there's a concept of phenomenal redness whose possession requires the ability to imagine red, i.e. an ability-requiring concept of phenomenal redness. Then couldn't we also say that there is a concept of dog whose possession requires the ability to imagine dogs, i.e. an ability-requiring concept of dog? And suppose we say that there is a concept of phenomenal redness whose exercise requires having an experience as of red, i.e. an experience-involving concept of phenomenal redness. Then couldn't we also say that there is a concept of pencil whose exercise requires holding a pencil in one's right hand, i.e. a pencil-involving concept of pencil? The crucial observation here is that if we adopt very fine-grained criteria of individuation for phenomenal concepts (as the 'Fregean' account urges us to do), it would be unprincipled not to adopt equally fine-grained criteria for other families of concepts. There is nothing wrong with doing this, of course. The problem is simply that we end up with a picture on which the acquisition of experiential knowledge does not count at all as an extraordinary or significant epistemic progress: the epistemic progress that Mary makes with respect to phenomenal redness when she leaves the black-and-white room will not be too different from the epistemic progress that other subjects make with respect to dogs when they learn how to visually imagine dogs or with respect to pencils when they start holding a pencil in their right hand.

(A defender of the 'Fregean' approach could reply that adopting more fine-grained individuation criteria for one family of concepts and less fine-grained individuation criteria for another family of concepts need not be an unprincipled move. After all, we have the intuition that experiencing phenomenal redness teaches us something about phenomenal

122The point is noted also by Ball (2009, 953).

redness, whereas holding a pencil in one's right hand need not teach us anything about pencils. Hence, we have independent reasons for positing experience-involving phenomenal concepts, but not pencil-involving pencil concepts. But this reply gets the explanatory order backwards: on the 'Fregean' approach, the existence of experience-involving phenomenal concepts should explain and not be explained by the fact that experiencing phenomenal redness teaches us something about it. So whatever reasons there might be not to extend the 'Fregean' approach to pencil concepts, the observation that holding a pencil in one's right doesn't teach us anything about pencils should not be one of them).

Intuitive worries – An even deeper, even if slightly more elusive worry, is the following. The 'Fregean' approach is committed to saying that what the inexperienced gain when they cease to be inexperienced is just a new way of representing or conceptualizing facts that they may have already been able to represent or conceptualize in other ways. For example, Mary may gain knowledge of the proposition *that red objects cause human beings to instantiate phenomenal redness**, but she already knew *that red objects cause human beings to instantiate phenomenal redness*: the two propositions have different concepts as constituents, but they target exactly the same fact – they are made true by the same portion of the world, as it were. We could say that, according to the 'Fregean' approach, Mary doesn't gain knowledge of any new fact, but just new knowledge of an old fact (i.e. new knowledge of a fact that she already had knowledge of, albeit under different modes of presentation).

Now, bracketing for a moment the worries I voiced above, it may be perfectly legitimate to model phenomenal knowledge in this way, i.e. as just another Frege-case along the lines of more famous cases involving Hesperus and Phosphorus or Superman and Clark Kent. And yet this seems to run against our intuitions. Intuitively, Mary does not gain new knowledge of a fact that she already had knowledge of under other modes of presentation.

Intuitively, the fact that she gains knowledge of when she comes to know what it is like to see red is a *purely experiential* fact. And when a fact is purely experiential no intelligible distinction can be drawn between knowing it 'experientially' and knowing it at all: knowing it 'experientially' *just is* knowing it, period. So, intuitively, no matter how many facts Mary knew when she was inside the black-and-white room, the knowledge she gains upon her release is *not* new knowledge of any of those facts – it is knowledge of a *brand new fact*.

Though certainly less than clear, these intuitions have undeniable allure. Perhaps one way of bringing them into sharper focus is to note that there is a contrast between what we are inclined to say about cases of acquisition of experiential knowledge and more familiar Frege-cases. It is not too much of a stretch to say that, if I know *that Clark Kent entered the bank*, I 'somehow' know that Superman entered the bank. By contrast, it seems completely false to say that, by knowing any of the things she knew when she was inside the black-and-white room, Mary 'somehow' knew what it is like to see red.[123] One would like to say that experiential knowledge was *completely* beyond reach for Mary until her release. In this sense, the 'Fregean' approach concedes far too much when it concedes that the subject-matter of phenomenal knowledge could have been known by Mary, *under some guise or another*, before her release.

123In principle, one could explain away this intuition in the following way. Someone who knows *that Clark Kent entered the bank* 'somehow' knows that Superman entered the bank in the sense that he knows of someone who is in fact Superman that he entered the bank (i.e. he has *de re* knowledge, concerning Superman, to the effect that he entered the bank). Since Mary has no *de re* knowledge of phenomenal redness, her case is different from that of someone who knows *that Clark Kent entered the bank*, and that's pretty much all we need to explain the asymmetry.

But this strategy fails, for two reasons. The first is that Mary has (or can be taken to have) *de re* knowledge of phenomenal redness (recall: there's no good reason to deny that her thoughts could 'latch onto' various phenomenal properties, including phenomenal redness). The second is that the intuition would arise even in the case of someone who does not have *de re* knowledge, concerning Superman, to the effect that he entered the bank (for example someone who knows only *that whoever actually crossed the street a minute ago entered the bank*).

I am prepared to concede that the considerations I've offered so far do not constitute conclusive evidence that the 'Fregean' approach to *Inexperience* is inadequate. The 'Fregean' approach has theoretical costs, but there's hardly any view in philosophy that doesn't come with a price. The approach runs the risk of trivializing the epistemic progress one makes when one acquires experiential knowledge, but maybe it's a good thing to dispel the air of extraordinariness that surrounds experiential knowledge. Finally, the consequences of the approach might be somewhat counterintuitive, but the intuitions it runs against are quite obscure and, in any case, intuitions as such are not sacred. Still, I think that the worries I've explored should motivate a certain more-than-superficial dissatisfaction with the 'Fregean' approach. And since the 'Fregean' approach is, arguably, the best approach to *Inexperience* that someone who embraces MVM can come up with, the worries I've explored invite us to look beyond MVM and its take on experiential knowledge.

3. *Inexperience* and the Subjectivist View of the Mental

MVM is an objectivist view – a view according to which the world is objectively the way it is. So someone who acceptsMVM has no choice but to treat experiential truths as objective truths. That is to say, someone who accepts MVM has no choice but to accept the principle I called *Objectivity* at the beginning:

(*Objectivity*) Every experiential truth is an objective truth.

It is because they accept *Objectivity* that friends of MVM have to explain the truth of

Inexperience in epistemological terms. If experiential truths are objective truths – truths that hold even from the point of view of the inexperienced – the reason why the inexperienced cannot know them must lie in the nature of experiential knowledge. One way or another, experiential knowledge must be tied with the fulfilment of requirements that the inexperienced, because of the condition they find themselves in, do not fulfil.

In the last section, I raised some difficulties for the project of explaining *Inexperience* in this way. A non-epistemological approach to *Inexperience* becomes possible if we reject *Objectivity*, viz. if we start thinking of experiential truths as subjective truths - truths that are *not* there for all of us and, therefore, do *not* form part of the true and complete story about how the world objectively is.

What we need in order to get the account going are three things. The first is the hypothesis that the propositions knowledge of which is experiential knowledge are subjective (i.e. that they can be true from some points of view and not others):

(*Subjectivity*) Every experiential proposition is subjective.

The second is the hypothesis that the subjects we characterize as 'inexperienced' are subjects from whose point of view certain experiential propositions fail to be true. More precisely:

(*Subjective Inexperience*) To be inexperienced is to be a subject from whose point of view the propositions knowledge of which is knowledge of what it is like to undergo a certain kind of experience are not true.

And the third is the principle of *Relative Factivity* which is already part of SVM:

(*Relative Factivity*) For any subject x, if x knows a proposition p then p is true from x's point of view.[124]

These three ingredients provide a neat explanation of why *Inexperience* holds. *Relative Factivity* says that if you know a certain proposition, that proposition must be true from your point of view. If it is assumed that this is necessarily the case (as seems plausible), *Relative Factivity* and *Subjective Inexperience* together entail that, necessarily, if one is inexperienced one lacks knowledge of what it is like to undergo a certain kind of experience. In principle this would still be compatible with at least certain kinds of experiential knowledge having no associated condition of inexperience (in the sense specified by *Subjective Inexperience*). But *Subjectivity* says that *all* experiential propositions are subjective (i.e. that each of them could be true from some points of view and not others) and experiential knowledge is nothing else than knowledge of experiential propositions (at least, if *Propositionality* is true). So the prediction is that, for every kind of experience, there will be actual or possible conditions of inexperience making it impossible to know what it is like to undergo an experience of that kind. And this is just what *Inexperience* says, at least in its most general formulation.

To be absolutely clear on this point: on the account we are considering, the reason why inexperienced subjects lack experiential knowledge is *not* that they lack a certain type of evidence or that they cannot entertain the propositions knowledge of which is experiential knowledge. The reason is, rather, that experiential truths *are not part of the world they live in*: certain experiential truths are not *truths* from their point of view. For example, when I know what it is like to see red, I know something which is not true (and, therefore, is not there to be known) for someone like Frank Jackson's Mary. Symmetrically, anyone who knows what it is like to taste Vegemite knows something which is not true (and, therefore, is not there to be

124See Chapter 3, § 3.

known) from my point of view.[125]

So far, I haven't said anything about the propositions knowledge of which is experiential knowledge, except that they are subjective propositions (i.e. propositions that could be true from some points of view and not others). Later on, I will make some specific suggestions as to what these propositions might be and I will explain how their subjective truth (or falsehood) makes for certain subjects being experienced (or inexperienced) in the particular way described in § 1. But before I do that, let me briefly review the advantages of the subjectivist account I am putting forward as a general explanatory strategy, in abstraction from any particular way of filling in the details.

The first thing to notice about the subjectivist account of *Inexperience* is that it allows us to keep the epistemology of experiential truths plain and simple: we don't have to impose any unprincipled restriction on the possible sources of experiential evidence, we can be as liberal as we want about what acquaintance with a property requires (in fact, we can even decide to get rid of acquaintance altogether) and we can, if we want, do away with at least some of the unusual concepts and modes of presentation that the Fregean approach commits us to (I am thinking, in particular, of strange beasts like experience-involving concepts). Now,

125It is interesting to notice that this way of explaining *Inexperience* is just an instance of a more general style of explanation, one that finds application in other cases involving factive propositional attitudes. Consider, for example, the attitude expressed by '…is relieved that p': one cannot be relieved that p unless p is the case. Arthur Prior famously complained that the truth I find relief in when I say 'Thank goodness that's over!' cannot be sempitenal: if it were, I could have found relief in it well before the end of the painful event (Prior 1959). The general pattern of explanation is as follows: if truth p were sempiternal, it wouldn't be rational of subject x (e.g. my earlier self) not to find relief in it, but certainly it is rational of subject x not to be relieved, so p can't be sempiternal. In the case of experiential truths, the reasoning takes the following form: if experiential truth q were objective, sooner or later some inexperienced individual y would certainly discover q (while remaining inexperienced); but we know that no inexperienced individual will ever discover q (while remaining inexperienced), so q can't be an objective truth, after all. Someone might ask whether, for this explanation of *Inexperience* to succeed, experiential truths have to be regarded as temporary truths. The answer is 'no'. More about this in the next section.

if keeping the epistemology of experiential truths plain and simple were the only reason to endorse the subjectivist account, one could reasonably argue that the account is too costly for what it delivers. But the nice thing about the subjectivist account is that it rests on hypotheses that (as I showed in Chapter 1) find independent support in our pre-theoretic intuitions and that (as I argued in Chapter 3 and 4) can help us solve also other puzzles concerning the epistemology of the mental. Moreover, the very idea that we need subjective facts and subjective truths to explain the peculiar nature of experiential knowledge is far from counterintuitive and has been in the air for a long time. For example, talking about the difference between knowledge of experiential truths and knowledge of non-experiential ones, Nagel writes:

> In the case of experience, […] the connection with a particular point of view seems much closer. It is difficult to understand what could be meant by the *objective* character of an experience, apart from the particular point of view from which its subjects apprehends it. After all, what would be left of what it was like to be a bat if one removed the viewpoint of the bat? (Nagel 1974, 443)

And talking about Frank Jackson's argument from *Unscrutability* to the falsity of *Physicalism*, Crane writes:

> Frank Jackson's famous 'knowledge argument' [...] is designed to show, from apparently uncontroversial premises and simple reasoning, that the physicalist conception of the world is false. [But it has been pointed out] that if sound, the argument would show more than that: it would show that some facts are subjective, and thus that a view which says that all facts are objective would be false. (Crane 2003, 69)

(Notice that I am not saying that Nagel means by "point of view" what I mean by "point of

view", nor that Crane means by "subjective facts" what I mean with that expression. They do not. All I am saying is that there is a very tempting idea that Nagel and Crane are voicing – the idea that we need subjective facts or subjective truths to explain the irreducibly subjective nature of experiential knowledge – and that the subjectivist account defended here represents one independently motivated way of glossing that idea).

Another very general point concerns the kind of explanation of *Inexperience* that the subjectivist account proposes. Recall what *Inexperience* says: for every kind of experience, there are distinctive conditions of inexperience such that, necessarily, anyone who is in those conditions does not know what it is like to undergo an experience of that kind. For a subjectivist the *explanans* of this is, quite simply, that experiential truths *are not truths* from the point of view of the inexperienced. This strikes me as a more basic and thoroughgoing *explanans* than any of those the objectivist can come up with. For example, when one is told that the inexperienced lack the evidence one needs in order to attain experiential knowledge, one is still strongly tempted to ask 'Why do they lack it?'. Similarly, when one is told that the inexperienced are not acquainted with certain phenomenal properties or that they lack certain concepts referring to those properties, one is still tempted to ask 'Why is it so?'. In other words, the various hypotheses in terms of which an objectivist can try to explain *Inexperience* do not seem to be natural *endpoints* for a philosophical explanation. By contrast, the claim that the propositions knowledge of which is experiential knowledge are not true from the point of view of the inexperienced does not seem to call for further explanation: if asked why those propositions are not true from the point of view of the inexperienced, we can naturally reply that *that's what it is* for the subjects in question to be inexperienced – the world they live in is one in which certain experiential facts simply do not obtain.

Other advantages of the subjectivist account become apparent when we compare it with the Fregean approach (which, I've suggested, represent the most promising approach that

an advocate of MVM could adopt). It's clear that someone who embraces the subjectivist account does not trivialize the nature of the epistemic progress involved in the acquisition of experiential knowledge. One can formulate very specific hypotheses about the propositions knowledge of which is experiential knowledge (two such hypotheses will be offered below and none of them entails that part of what Mary comes to know or is in a position to come to know when she leaves the black-and-white room is *that phenomenal redness* is phenomenal redness**). Moreover, given the plausible assumption that only a tiny portion of the facts we know or are in a position to know every day are subjective facts, the subjectivist account respects the idea that ceasing to be inexperienced allows one to learn something in a way in which acquiring the ability to imagine dogs or starting to hold a pencil in one's right hand do not.

It's also clear that someone who adopts the subjectivist account is better-placed to accommodate certain intuitions about experiential knowledge than someone who adopts the Fregean approach. Recall that, according to the Fregean approach, there is a sense in which inexperienced subjects *can* attain knowledge of whatever facts, out there in the world, constitute the subject matter of phenomenal knowledge – the point is just that they cannot attain knowledge of those facts *under all modes of presentations*. However, according to our intuitions, the facts one comes to know when one comes to know what it is like to undergo this or that experience are *fully experiential* facts, and when a fact is fully experiential no intelligible distinction can be drawn between knowing it experientially and knowing it at all. Subjectivists can do justice to this idea. According to them, before being exposed to an instance of red, Mary couldn't even *start* to know the fact one knows when one knows what it is like to see red: from Mary's perspective, that fact was just not there to known, *under any mode of presentation whatsoever*, until the moment of her release.

Let me now move on to say something – for the most part, in an exploratory and

speculative spirit – about the propositions knowledge of which is experiential knowledge. In the rest of this section, I will put forward some specific proposals as to the nature of these propositions. I will assume the correctness of SVM. And I will also also assume that, for every kind of experience, there is no unique proposition p such that knowing what it is like to undergo an experience of that kind requires knowing p (in other words, I will assume the falsity of the principle I called *Uniqueness* in § 1). My focus will be on the case of knowing what it is like to see red, but it is not difficult to see how hypotheses similar to the ones I will discuss can be advanced, mutatis mutandis, for other types of experiential knowledge.

In answering the question 'What kind of propositions is knowledge of what it is like to be red knowledge of?', a subjectivist who is interested in explaining *Inexperience* along the lines suggested above should keep several constraints in mind. The first and most obvious is that there must be some intuitive plausibility in the supposition that, by knowing one or another of the propositions in question, one knows what it is like to see red. The second is that (in line with *Subjectivity*) the propositions in question should not hold true from all points of view (or, at least, not necessarily so). Thirdly and relatedly, if one is to comply with *Subjective Inexperience*, one should choose propositions that can be taken to be false from the point of view of precisely those subjects who, intuitively, are in a condition of inexperience that prevents them from knowing what it is like to see red (if what I said in § 1 is correct, these will be subjects who lack the ability to imagine or recognize red and are not undergoing any experience as of red).

I think there are at least two kinds of propositions that fulfill these requirements. The first kind includes propositions like the following:[126]

126On some accounts of propositions, these may well count as the same proposition. I remain neutral on the matter.

there is an EXPERIENCE *as of red*

PHENOMENAL REDNESS is instantiated

it EXPERIENCES redly (with "it" used as in "It is snowing")

...

It seems plausible to think that by knowing any of these propositions one knows what it is like to see red. To illustrate this with a familiar example, consider again Conee's Mary, who lacks visual imagination, but comes to know what it is like to see red by intently gazing at the color of red ripe tomatoes. If SVM is correct, it doesn't seem at all unnatural to suppose that she does this by coming to know one of the propositions above (in fact, the supposition is all the more plausible given the 'Lichtenbergian' hypothesis defended in Chapter 4 about the content of the judgments we form about our own experiences). But notice the propositions above are not objectively true. In particular, the propositions above, while true from Mary's point of view (at least as long as she intently gazes at the color of red ripe tomatoes), will be false from the point of view of those who find themselves in a certain condition of inexperience (i.e. those who, among other things, are not having any experience as of red). In fact, if SVM is correct (in particular if the disjunctive definition of experience in terms of EXPERIENCE is correct), it can be argued that the very fact that a subject is not undergoing any experience as of red is grounded in the fact that there is no EXPERIENCE as of red from that subject's point of view. So the propositions above have the desired feature of being ones whose falsity from one's point of view is partly responsible for one's being inexperienced (in a certain way), just as *Subjective Inexperience* dictates.

Needless to say, propositions like *there is an* EXPERIENCE *as of red*, *PHENOMENAL REDNESS is instantiated* or *it EXPERIENCES redly* (with "it" used as in "It is snowing") can't be the *only* propositions knowledge of which is knowledge of what it is like to see red. Right

now, I am not intently gazing at any red object (nor having any experience as of red). So, from my point of view, there is *no* EXPERIENCE as of red, PHENOMENAL REDNESS is *not* instantiated and it does *not* EXPERIENCE redly (remember: from my point of view, only my experiences of red are EXPERIENCES of red). Yet I know what it is like to see red. What kind of propositional knowledge makes *me* knowledgeable of what it is like to see red? Maybe knowledge of one or another of the following propositions:[127]

there is a way of IMAGINING red

red is IMAGINABLE

it is IMAGINABLE that there is red (with "it" used as in "It is snowing")

...

From my point of view, even when I'm not looking at any red objects (or having any experience as of red), red is an intimately *familiar* property – it is IMAGINABLE (i.e. strikingly and quintessentially imaginable) and not just imaginable. The idea, then, is that it is in virtue of knowing that red is so intimately familiar that I know what it is like to see red. But of course this IMAGINABILITY is a subjective property of red: to say that red is IMAGINABLE (or that there is a way of IMAGINING red) is to say that red can easily be IMAGINED,[128] and this need not

127Again, I remain neutral on whether these should count as one and the same proposition.

128Admittedly, this analysis is only *roughly* correct. On one way of reading this expression, to say that red "can easily be IMAGINED" is just to say that, in some nearby worlds, red is being IMAGINED. But we don't want to say that, from Mary's point of view, red is IMAGINABLE (or that there is a way of IMAGINING red) whenever, in some nearby possible world, Mary is IMAGINING red (at least not in the sense of "IMAGINABLE" and "way of IMAGINING red" that are relevant here). After all, those nearby worlds might be ones in which Mary is outside the black-and-white room or has ingested red-experience-inducing pills: intuitively, the closeness of these worlds should not make red strikingly familiar to Mary in the actual world. What exactly might be involved in the IMAGINABILITY of red (or in the existence of a way of IMAGINING red) is a delicate question that I don't have any terribly informative answer to. (The situation is not entirely unlike that of other dispositional properties – fragility, for instance – that do not lend themselves to obviously true and informative analyses).

be the case from all points of view. In particular, red will *not* be IMAGINABLE from the point of view of those subjects who find themselves in a condition of inexperience (i.e. those who, among other things, lack the ability to imagine or recognize red). In fact, if SVM is correct (in particular if the disjunctive definition of imagining in terms of IMAGINING is correct) it might be argued that the fact that a subject does not have the ability to imagine red is grounded in the fact that red is not IMAGINABLE from that subject's point of view. So, once again, the propositions above have the desired feature of being ones whose falsity from one's point of view is partly responsible for one's being inexperienced (in a certain way), just as *Subjective Inexperience* requires.

The proposal that knowledge of what it is like to see red can be knowledge that there is a way of IMAGINING red resembles an account of experiential knowledge explored by Stanley and Williamson (2001), who suggested that "x's knowing how to imagine red amounts to knowing a proposition of the form 'w is a way for x to imagine red', entertained under a guise involving a practical mode of presentation of a way" (442). But there are two important differences.

First of all, a subjectivist doesn't have to bring in practical modes of presentation (in fact, faithful to the idea that concepts and modes of presentation can be possessed deferentially and, therefore, be shared across individuals with different skills and abilities, subjectivists can refuse to posit any such things as *practical* modes of presentation). Stanley and Williamson need practical modes of presentation for a very simple reason: on their view, the proposition *that w is a way for Giovanni to imagine red*, being an objective proposition, could be true (and actually be known by Giovanni under some conceptual guise or another) *before* Giovanni actually learns how to imagine red. For notice that w could be the way someone else than Giovanni is imagining red, so Giovanni could refer to that way of imagining red demonstratively, be told that that's a way for him to imagine red and thereby

learn that that's a way for him to imagine red. Not so on my proposal: since it is only Giovanni who has ways of IMAGINING things, no way of IMAGINING red exists until Giovanni acquires the ability to imagine red.

Another important difference is that a subjectivist doesn't have to build reference to Giovanni into the proposition Giovanni learns when he learns how to imagine red: that proposition is the perfectly general proposition *that there is a way of IMAGINING red* (or *that red is IMAGINABLE* or *that it is IMAGINABLE that there is red* with "it" used as in "It is snowing"). This ensures that, when two subjects learn what it is like to see red, there is a single proposition that they are both learning and a single kind of fact about the world that, from their respective points of view, they are getting hold of. This commonality, across different subjects, of the objects of experiential knowledge is a distinctive feature of the subjectivist proposal vis-à-vis at least some of its competitors and I will come back to it in the next section.

Notice that, even if the particular propositions I offered are propositions knowledge of which is knowledge of what it is like to see red, they need not be the only propositions having this property. In this sense, my account is somewhat open-ended: much more would have to be said about each of the hypotheses I've put forward and more alternatives would have to be explored and carefully evaluated. But the purpose of this section – and of the chapter as a whole – was not to say the last word about what it is that we know when we know what it is like to see red. It was, rather, to show that someone who embraces SVM has a natural and simple account to offer of *Inexperience* – an account that is, arguably, simpler and more natural than any account premised on the truth of *Objectivity*.

4. Objections

To conclude, I want to briefly consider three objections to the account of *Inexperience* I have been outlining as well as to the specific proposals I have put forward.

The first objection is that for the account of *Inexperience* I have been outlining to work we need more than just *Subjectivity*:

(*Subjectivity*) Every experiential proposition is subjective (i.e. it could be true from some points of view and not others)

We need also:

(*Temporariness*) Every experiential proposition is temporal (i.e. it could be true at some times and not others)

To see this, suppose that the propositions knowledge of which is knowledge of what it is like to see red were sempiternal, i.e. true at all times if true at all. Then the subjectivist would be left with no explanation of why subjects like Frank Jackson's Mary do not know them. For example, the proposition *that there is a way of* IMAGINING *red*, which is supposed to be true from Mary's point of view *after* she acquires the ability to imagine red, would be always true if true at all and so, in particular, it would be true from Mary's point of view even *before* Mary's release from the black-and-white room. But then the story I told in the last section about why Mary couldn't know that proposition before her release (a story involving *Relative Factivity* and the falsity of the proposition) would not work.

My reply to this objection is as follows. I agree that it's natural to think of experiential

propositions as *both* temporal *and* subjective (in fact, I think it's natural to think of them as *contingent*, temporary and subjective). But I do not agree that the story I told in the last section couldn't be made to work without assuming *Temporariness* (or, for that matter, some analogous principle saying that experiential propositions are contingent). What's essential to the story I told in the last section is the idea that the point of view of an inexperienced subject is a point of view from which certain experiential propositions are not true (that's, in essence, what *Subjective Inexperience* says). Now, if *Temporariness* is false, this idea is incompatible with (a) Mary being an inexperienced subject and (b) there being just *one* point of view corresponding to her entire mental life. For if there is just one point of view corresponding to Mary's entire life and, at some point of her life, Mary knows what it is like to see red, the proposition knowledge of which is knowledge of what it is like to see red has to be true from that point of view and so, by *Subjective Inexperience*, Mary can't be an inexperienced subject, after all. But if *Temporariness* is false, there is no reason to think that there is just *one* point of view corresponding to Mary's entire mental life. On the contrary, if *Temporariness* is false, we can think that there will be as many points of view as are the instantaneous stages of Mary's mental life: everything that is the case from each of these points of view will be sempiternally the case and, from each of these points of view, everything the case from other points of view will be otherpersonally the case (for each of Mary's instantaneous selves, past and future selves will be a bit like other minds). Then the rest of the story goes as usual: Mary's inexperienced stages lack experiential knowledge because the propositions knowledge of which is experiential knowledge are not true from their point of view. So it's not true that the story I told in the last section requires the truth of *Temporariness*.

The other two objections I want to consider are, really, ways of missing the point of the subjectivist account.

According to the first objection, what the subjectivist account says is, really, that

experiential knowledge is knowledge of some indexical proposition. When, for instance, the subjectivist says that knowledge of what it is like to see red can be knowledge *that* PHENOMENAL REDNESS *is instantiated*, what she is really saying is that knowledge of what it is like to see red can be knowledge *that one instantiates phenomenal redness* – knowledge one would express using the sentence "*I* instantiate phenomenal redness". One doesn't need to embrace Subjectivism or draw any distinction between phenomenal redness and PHENOMENAL REDNESS to model experiential knowledge in this way: all one needs is a proper understanding of how indexical expressions and indexical concepts work.

The objection overlooks one crucial difference between the account I'm proposing and the 'indexical' account. On the account I'm proposing, when you and I both learn what it is like to see red in the experience-requiring sense, we both acquire knowledge *of the very same proposition*. This is the simple and perfectly general proposition *that* PHENOMENAL REDNESS *is instantiated.* It seems to me that this is just how it should be: knowledge of what it is like to see red is not (or not primarily) knowledge that I have *concerning myself* and you have *concerning yourself*, but knowledge concerning the experience of seeing something red. Not so on the indexical account. On the indexical account sketched above, there is no single proposition that both you and I know when we know what it is like to be seeing red: you know that *you* instantiate phenomenal redness and I know that *I* instantiate phenomenal redness. Two different propositions are at stake here, notwithstanding the fact that we could express them using the same words, i.e. "*I* instantiate phenomenal redness". Whatever its merits, this 'indexical' account is *not* the one I've been defending (it is, in fact, just a particular version of the 'Fregean' approach).

According to the second objection, we could achieve the same results of the subjectivist account without embracing Subjectivism, but only some version of Subject-relativism. Subjectivism is the doctrine that some propositions are true (i.e. true absolutely or

simpliciter) without being true from every subjective point of view. Subject-relativism, on the other hand, is the more modest claim that some propositions are true relative to some points of view and not others (without any commitment to their being true absolutely or simpliciter).[129] A subject-relativist could agree that you and I both know *the same proposition* when we acquire occurrent knowledge of phenomenal redness. They could also agree that that proposition is the proposition *that PHENOMENAL REDNESS is instantiated*. But they could eschew commitment to Subjectivism by insisting that that proposition is only true relative to this or that point of view, not true absolutely or simpliciter.

My response is that Relativism does *not* get us what full-blooded Subjectivism gets us. Subjectivism gets us the idea that, when you and I both know what it is like to see something red, there is a certain kind of fact that my knowledge targets and, from your point of view, it is *exactly the same kind of fact* that your knowledge targets (i.e. the kind of fact that, from my point of view, could be expressed by the proposition, true absolutely and simpliciter, *that there is PHENOMENAL REDNESS*). It seems to me that this is just how it should be: intuitively, my knowledge of what it is like to see red puts exactly the same demands on reality that, from your point of view, your knowledge of what it is like to see red puts on reality. Not so on a relativist account on which the proposition *that PHENOMENAL REDNESS is instantiated* is only true relative to this or that point of view. On that account, my knowledge and your knowledge are essentially involved with two different points of view: the fact targeted by *my* knowledge of what it is like to see red is the fact that the proposition *that PHENOMENAL REDNESS is instantiated* is true relative to *my* point of view, whereas the fact targeted by *your* knowledge of what it is like to see red is the fact that the same proposition is true relative to *your* point of view. Whatever its merits, a subject-relativist account of this sort is *not* the one I've been defending (it is, in fact, just another version of the 'Fregean' approach).

129The difference between Subjectivism and subject-relativism was discussed in Chapter 1, § 1.

References

Alexander, Samuel. (1920). *Space, Time, and Deity*, vol. 2. London: Macmillan.

Alston, William. (1971). "Varieties of Privileged Access". *Americal Philosophical Quarterly* 8 (3): 223-241.

Alter, Torin. (2001). "Know-how, Ability, and the Ability Hypothesis". *Theoria* 67: 229-39.

Anscombe, Elizabeth. (1975). "The First Person" in Guttenplan, S. (Ed.), *Mind and Language*, Oxford: Clarendon Press (pp. 45-65).

Armstrong, David M. (1968). *A Materialist Theory of the Mind*, London: Routledge.

Armstrong, David M. (1981). *The Nature of Mind and Other Essays*, Ithaca, NY: Cornell University Press.

Augustine (1974). "De Trinitate" in Schaff, P. (Ed.), *Nicene and Post-Nicene Fathers of the Christian Church*, vol. Ill, Grand Rapids: Eerdmans.

Avramides, Anita. (2001). *Other Minds*. London: Routledge.

Ayer, Alfred J. (1990). *The Problem of Knowledge*. London: Penguin Books.

Ball, Derek. (2009). "There Are No Phenomenal Concepts". *Mind* 118 (472): 935-962.

Baron-Cohen, Simon.(1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

Bayne, Tim and David J. Chalmers. (2003). "The Unity of Consciousness" in Cleeremans, A. (Ed.), *The Unity of Consciousness: Binding, Integration and Dissociation*, Oxford:

Oxford University Press.

Berker, Selim. (2008). “Luminosity Regained”. *Philosophical Imprint* 8: 1-22.

Bilgrami, Akeel. (2012). “The Unique Status of Self-knowledge” in Coliva, A. (Ed.), *The Self and Self-knowledge*, Oxford: Oxford University Press.

Biro, John I. (1991). “Consciousness and Subjectivity”. *Philosophical Issues*, 1: 113-133.

Block, Ned. (1978). “Troubles with Functionalism”. *Minnesota Studies in The Philosophy of Science* 9: 261–325.

Block, Ned. (2007). “Max Black’s Objection to Mind-Body Identity” in *Phenomenal Concepts and Phenomenal Knowledge*, ed. Torin Alter and Sven Walter, Oxford: Oxford University Press (pp. 248-306).

Block, Ned; Flanagan, Owen and Güven Güzeldere (Eds.). (1997). *The Nature of Consciousness: Philosophical Debates*. Cambridge (MA): MIT Press.

Block, Ned and Robert Stalnaker. (1999). “Conceptual Analysis, Dualism, and the Explanatory Gap”. *Philosophical Review* 108: 497-528.

Bolzano, Bernard. (1972). *Theory of Science*, edited and translate by R. George. Berkeley: University of California Press.

Broad, Charlie D. (1923). *Scientific Thought*. New York: Harcourt, Brace.

Brogaard, Berit. (2007). “Span Operators”. *Analysis* 67 (1): 72-79.

Brogaard, Berit. (2012). *Transient Truth: An Essay in the Metaphysics of Propositions*. New York: Oxford University Press.

Brower-Toland, Susan. (2012). “Medieval Approaches to Consciousness: Ockham and Chatton”. *Philosophers' Imprint* 12 (17): 1-29.

Burge, Tyler. (1979). “Individualism and the Mental”. *Midwest Studies in Philosophy.*

Burge, Tyler. (1986). “Intellectual Norms and the Foundations of Mind”. *Journal of Philosophy*, 83: 697–720.

Cameron, Ross. (2008). “Turtles all the way down: regress, priority and fundamentality”. *The Philosophical Quarterly* 58, 230: 1-14.

Campbell, John. (1999). “Immunity To Error Through Misidentification And The Meaning Of A Referring Term”. *Philosophical Topics* 26 (1-2): 89-104.

Cappelen, Herman and John Hawthorne. (2009). *Relativism and Monadic Truth*. Oxford: Oxford University Press.

Cappelen, Herman and John Hawthorne. (2011). “Reply to Lasersohn, MacFarlane, and Richard”. *Philosophical Studies* (156): 449-466.

Cassam, Quassim. (1999). *Self and World*. Oxford: Oxford University Press.

Cassam, Quassim. (2004). “Introspection, Perception and Epistemic Priviledge”. *The Monist* 87 (2): 255-274.

Cassam, Quassim. (2007a). “Ways of Knowing”. *Proceedings of the Aristotelian Society* 107 (3): 339-358.

Cassam, Quassim (2007b). *The Possibility of Knowledge*. New York: Oxford University Press.

Cary, Philipp. (2000). *Augustine's Invention of the Inner Self: The Legacy of a Christian*

Platonist. Oxford: Oxford University Press.

Chalmers, David J. (1996). *The Conscious Mind. In Search For A Fundamental Theory*. New York: Oxford University Press.

Chalmers, David J. (2002a). "Does Conceivability Entail Possibility?" in Gendler, T.S. and J. Hawthorne (Eds.), *Conceivability and Possibility*, Oxford: Oxford University Press.

Chalmers, David J. (2002b). "Phenomenal Concepts and the Knowledge Argument" in Ludlow, P., Nagasawa, Y. and D. Stoljar (Eds.), *There's Something About Mary*, Cambridge (MA): MIT Press.

Chalmers, David J. (2003a). *The Character of Consciousness*. Oxford: Oxford University Press.

Chalmers, David J. (2003b). "Consciousness and its place in nature" in Stitch, S. And T. Warfield, *The Blackwell Guide to Philosophy of Mind*, Oxford: Blackwell Publishing.

Chalmers, David J. (2012). *Constructing the World*. Oxford: Oxford University Press.

Chalmers, David J. and Frank Jackson. (2001). "Conceptual Analysis and Reductive Explanation". *Philosophical Review* 110, 3: 315-360.

Churchland, Paul S. (1983). "Consciousness: the Transmutation of a Concept". *Pacific Philosophical Quarterly*, 64: 80-95.

Cohen, Stewart. (1986) ''Knowledge and Context''. *Journal of Philosophy* 83 (10): 574–583.

Coliva, Annalisa. (2006). "Error Through Misidentification: Some Varieties". *The Journal of Philosophy* 103 (8): 403-425

Conee, Earl. (1994). “Phenomenal Knowledge”. *Australasian Journal of Philosophy* 72: 136-50.

Conee, Earl. (2005). “The Comforts of Home”. *Philosophy and Phenomenological Research* 70 (2): 444-451.

Correia, Fabrice and Sven Rosenkranz. (2011). *As Time Goes By*. Paderborn: Mentis.

Correia, Fabrice and Sven Rosenkranz. (2012). “Eternal Facts in an Ageing Universe”. *Australasian Journal of Philosophy* 90 (2): 307-320.

Correia, Fabrice and Benjamin Schnieder. (Eds.). (2012). *Metaphysical Grounding*, Cambridge: CUP.

Craig, William Lane. (1997). “Is Presentness a Property?”. American Philosophical Quarterly 34, 1: 27-40.

Crane, Tim. (2003). “Subjective Facts” in Lillehammer, H. and G. Rodriguez-Pereyra, *Real Metaphysics*, London: Routledge.

Crisp, Thomas M. (2004). “On Presentism and Triviality” in Zimmerman, Dean (Ed.), *Oxford Studies in Metaphysics (Vol. 1)*, Oxford: Clarendon Press.

Crisp, Thomas M. (2005). “Presentism and ‘Cross-time’ Relations”. *American Philosophical Quarterly*, 42 (1): 5-17.

Davidson, Donald. (1973). “Radical Interpretation”. *Dialectica* 27: 314–28.

Davidson, Donald. (1983). “A Coherence Theory of Truth and Knowledge” in Henrich, D. (Ed.), *Kant oder Hegel?*, Stuttgart: Klett-Cotta.

De Clercq, Rafael. (2006). “Presentism and the Problem of Cross-time Relations”.

Philosophy and Phenomenological Research, 72 (2):386-402.

Dennett, Daniel. (1991). *Consciousness Explained*. Boston: Little, Brown and Company.

De Rose, Keith. (1992) ‘‘Contextualism and Knowledge Attributions’’. *Philosophy and Phenomenological Research* 52 (4): 913-929.

Descartes, René. (1988). *Selected Philosophical Writings*, translated by Cottingham, J., Stoothoff R. and D. Murdoch. Cambridge: Cambridge University Press.

De Sousa, Ronald. (2004). “Twelve varieties of subjectivity: dividing in hopes of conquest” in Larrazabal, J.M and L.A. Pérez Miranda (Eds.), *Language, Knowledge and Representation: Proceedings of the Sixth International Colloquium on Cognitive Science*, London: Kluwer Academic.

Douven, Igor. (2011). “Abduction”. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (Ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/abduction/>.

Dorr, Cian. (2010). “How To Be A Modal Realist”. (MS).

Dorr, Cian and Jeremy Goodman. (MS). “Diamonds Are Forever”

Egan, Andy. (2006). “Secondary Properties and Self-Location”. *Philosophy and Phenomenological Research* 72 (1): 97-119.

Egan, Andy (2007) “Epistemic Modals, Relativism and Assertion”. *Philosophical Studies* 133 (1): 1–22.

Egan, Andy, Hawthorne, John and Brian Weatherson (2005). “Epistemic Modals in Context” in Preyer, G. and P. Peter (Eds.), *Contextualism in Philosophy*, Oxford: Oxford University Press.

Evans, G. (1982). *The Varieties of Reference*. Oxford: Clarendon Press.

Feigl, Herbert. (1967). *The 'mental' and the 'physical': the essay and a postscript*. Minneapolis: University of Minnesota Press.

Fine, Kit. (2001). "The Question Of Realism". *Philosophers' Imprint* 1, 1: 1-30.

Fine, Kit. (2005). "Tense and Reality" in Kit Fine, *Modality and Tense: Philosophical Papers*, Oxford: Clarendon Press: 261-320.

Fine, Kit. (2006). "The Reality of Tense". *Synthese* 150: 399-414.

Fodor, Jerry. (1974). "Special sciences". *Synthese*, 28: 77-115.

Foster John. (1996). *The Immaterial Self: A Defence of the Cartesian Dualist Conception of Mind.* London: Routledge.

Frege, Gottlob. (1984). "Thoughts". In G. Frege, *Collected Papers on Mathematics, Logic, and Philosophy*, Oxford: Basil Blackwell, pp. 351–372.

Fricker, Elizabeth. (1998). "Self-knowledge: Special Access Versus Artefact of Grammar – A Dichotomy Rejected" in Wright, C., Smith, B. C. and C. Macdonald (Eds.), *Knowing Our Own Minds*, Oxford: Oxford University Press (pp. 155-206).

Garcia-Carpintero, Manuel. (forthcoming). "Self-Conception: Sosa on De Se Thought" in Turri, J. (Ed.), *Virtuous Thoughts: The Philosophy of Ernest Sosa*, Philosophical Studies Series 119.

Gertler, Brie. (2001). "Introspecting phenomenal states". *Philosophy and Phenomenological Research* 63: 305–328.

Gertler, Brie. (2011). "Self-Knowledge", *The Stanford Encyclopedia of Philosophy*,

Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/self-knowledge/>.

Glanzberg, Michael. (2009). “Semantics and Truth Relative to a World”. *Synthese*, 166: 281–307.

Goldman, Alvin I. (2006). *Simulating Minds: the Philosophy, Psychology, and Neuroscience of Mindreading*. New York: Oxford University Press.

Gopnik, Alison, Meltzoff, Andrew and Patricia Kuhl. (1999). *The Scientist in the Crib*. New York: HarperCollins.

Gopnik, Alison and Henry M. Wellman. (1992). “Why the child’s theory of mind really is a theory”. *Mind and Language* 7: 145–171.

Gopnik, Alison and Henry M. Wellman. (1994). “The Theory Theory” in Hirschfeld, L. and S. Gelman (Eds.), Mapping the Mind: Domain Specificity in Cognition and Culture. New York: Cambridge University Press.

Greenough, Patrick. (2012) “Discrimination and Self-Knowledge” Smithies, D. and D. Stoljar (Eds.), *Introspection and Consciousness*, Oxford: Oxford University Press.

Güzeldere, Güven. (1997). “The Many Faces of Consciousnee: A Field Guide” in Block et al. (1997).

Hare, Caspar J. (2009). *On Myself, and Other Less Important Subjects*. Woodstock: Princeton University Press.

Harré, Rom. (1999). “Nagel’s Challenge and the Mind-Body Problem”. *Philosophy* 74, 288: 247-270.

Hatfield, Gary. (2003). *Descartes and the Meditations.* London: Routledge.

Hawthorne, John P. (2002). “Blocking Definitions of Materialism”. *Philosophical Studies* 110: 103-113.

Hawthorne, John and David Manley. (2012). *The Reference Book*. Oxford: Oxford University Press.

Heidegger, Martin. (2001). *Being and Time*. Translated by Macquarrie, J. and E. Robinson. Oxford: Blackwell.

Hellie, Benj. (2004). “Inexpressible Truths and the Allure of the Knowledge Argument” in Nagasawa, Y., Ludlow, P. and D. Stoljar (Eds.), *There's Something About Mary*, Cambridge, MA: MIT Press. (pp. 333-)

Hellie, Benj. (2013). “Against Egalitarianism”. Analysis 73 (2): 304-320.

Hellie, Benj. (MS). *Out Of This World.*

Hill, Christopher. (1991). *Sensations: A Defense of Type Materialism*. Cambridge: Cambridge University Press.

Horgan, Terence and Uriah Kriegel. (2007). “Phenomenal Epistemology: What Is Consciousness That We May Know It So Well?”. *Philosophical Issues* 17 (1): 123–144.

Hossack, Keith. (2006). “Vagueness and Personal Identity” in MacBride, F. (Ed.), *Identity and Modality*, Oxford: Oxford University Press (pp. 221-241).

Horwich, Paul. (1998). *Truth*. New York: Oxford University Press.

Hofweber, Thomas. (2009). “Ambitious, yet modest, metaphysics” in Chalmers, D., Manley, D. Manley and R. Wasserman (Eds.), *Metametaphysics*, Oxford: Oxford

University Press.

Hume, David. (1978). *A Treatise of Human Nature*. Edited by Selby-Bigge, L.A. and revised by P.H. Nidditch. Oxford: Oxford University Press.

Husserl, Edmund. (1988). *Cartesian Meditations. An Introduction to Phenomenology*. Dordrecht: Kluwer Academic Publishers.

Hyslop, Alec. "Other Minds". *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (Ed.), URL = <http://plato.stanford.edu/archives/spr2014/entries/other-minds/>.

Jackson, Frank. (1982). "Epiphenomenal Qualia". *Philosophical Quarterly* 32: 127-136.

Jackson, Frank. (1994). "Armchair Metaphysics" in O'Leary-Hawthorne, J. and Michaelis. M. (Eds.), *Philosophy in Mind*, Dordrecht: Kluwer.

James, William. (1950). *Principles of Psychology*. New York: Dover Publications.

Johnston, Mark. (2009). *Introduction* to in Hare, C. J., *On Myself, and Other Less Important Subjects*, Woodstock: Princeton University Press (pp. xi-xviii).

Kant, Immanuel. (1998). *Critique of Pure Reason*. Translated by Guyer, P. and A.W. Wood. New York: Cambridge University Press.

Kant, Immanuel. (2007). *Critique of Judgment*. Translated by Meredith, J.C. New York: Oxford University Press.

Kaplan, David. (1989). "Demonstratives" in Almog, J., Perry, J. and H. Wettstein (Eds.), *Themes From Kaplan*, New York: Oxford University Press (pp. 481–563).

Kenny, Anthony. (1993). *Aquinas on Mind*. London: Routledge.

Kim, Jaegwon. (1993). *Supervenience and Mind*. Cambridge: Cambridge University Press.

Kim, Jaegwon. (1998). *Mind in a Physical World*. Cambridge, MA: MIT Press.

Kim, Jaegwon. (2003). "Supervenience, Emergence, Realization, Reduction" in Loux, M.J. and D.W. Zimmerman (Eds.), *The Oxford Handbook of Metaphysics*, Oxford: Oxford University Press.

King, Jeffrey. (2003). "Tense, modality, and semantic values". *Philosophical Perspectives: Language and Philosophical Linguistics*, 17: 195–245.

Kölbel, Max. (2003). "Faultless Disagreement". *Proceedings of the Aristotelian Society* 104: 53-73.

Lackey, Jennifer. (2011). "Assertion and Isolated Secondhand Knowledge" in Brown, J. and H. Cappelen (Eds.), *Assertion*, Oxford: Oxford University Press (pp. 251-275).

Lasersohn, Peter (2005). "Context Dependence, Disagreement and Predicates of Personal Taste". *Linguistics and Philosophy* 28: 643-686.

Leibniz, Gottfried W. (1989) *Philosophical Essays*. Eds. Ariew, R. and D. Garber. Indianapolis: Hackett.

Leng, Mary. (2010). *Mathematics and Reality.* Oxford: Oxford University Press.

Le Poidevin, Robin. (1998). "The Past, Present and Future of the Debate about Tense" in Le Podevin, R. (Ed.), *Questions of Time and Tense*, New York: Oxford University Press.

Leslie, Alan M. (1987). "Pretence and Representation: The Origins of ''Theory of Mind'". *Psychological Review* 94: 412-426.

Levin, Janet. (2007). "What Is A Phenomenal Concept?" in Alter, T. and S. Walter (Eds.),

Phenomenal Concepts and Phenomenal Knowledge, Oxford: Oxford University Press (pp. 86-110).

Levine, Joseph. (1983). "Materialism and qualia: the explanatory gap". *Pacific Philosophical Quarterly*, 64: 354-361.

Lewis, Clarence. I. (1964). "Some Logical Considerations Concerning the Mental" in Vesey, N. A. (Ed.), *Body and Mind*, London: George Allen and Unwin.

Lewis, David. (1966). "An Argument for The Identity Theory". *Journal of Philosophy* 63,1: 17-25.

Lewis, David. (1979). "Attitudes De Dicto and De Se". *Philosophical Review* 88: 513-543.

Lewis, David. (1983a). "Postscript to 'Mad Pain and Martian Pain'" in his *Philosophical Papers* Vol. 1., New York: Oxford University Press (pp. 130-132).

Lewis, David. (1983b). "Individuation by Acquaintance and by Stipulation". *Philosophical Review* 92 (1): 3-32.

Lewis, David. (1986) *On the Plurality of Worlds*. Oxford: Basil Blackwell.

Lewis, David. (1990). "What Experience Teaches" in William Lycan (Ed.), *Mind and Cognition: A Reader*, Oxford: Blackwell (pp. 499-519).

Lewis, David. (1996). "Elusive Knowledge". *Australasian Journal of Philosophy* 74 (4): 549-567.

Lichtenberg, Georg Christoph. (2012). *Philosophical Writings*. Translated and annotated by Tester, S. Albany, NY: State University of New York Press.

Loar, Brian. (1997). "Phenomenal States" (Revised Version) in Block, N., Flanagan, O. and

G. Guzeldere (Eds.), *The Nature of Consciousness*, Cambridge MA: MIT Press (pp. 597-616).

Locke, John. (1975). *An Essay Concerning Human Understanding*, edited by Peter. H. Nidditch. Oxford: Clarendon Press.

Ludlow, Peter. (2004). "Presentism, Triviality and the Variety of Tensism" in Zimmerman, Dean (Ed.), *Oxford Studies in Metaphysics (Vol. 1)*, Oxford: Clarendon Press.

Lycan, William G. (1990). "What is the 'subjectivity' of the mental". *Philosophical Perspectives*, vol. 4, Action Theory and Philosophy of Mind: 109-130.

MacFarlane, John. (2003). "Future Contingents and Relative Truth". *The Philosophical Quarterly* 53: 321-36.

MacFarlane, John. "Epistemic Modals Are Assessment-Sensitive" in Weatherson, B and A. Egan (Eds.), *Epistemic Modality*, Oxford: Oxford University Press.

Malebranche, Nicolas. (1997). *The Search After Truth*. Translated and edited by Lennon, T.M. and P.J. Olscamp, Cambridge: Cambridge University Press.

Markosian, Ned. (1993). "How Fast Does Time Pass?". *Philosophy and Phenomenological Research* 53: 829-844.

McDaniel, Kris. (2012). "Review of Caspar Hare's *On Myself, and Other, Less Important Subjects*". *Ethics* 122 (2): 403-410.

McDowell, John. (1982). "Criteria, Defeasibility and Knowledge". *Proceedings of the British Academy*, 68, 455-479.

McGeer, Victoria. (1996). "Is 'self-knowledge' an empirical problem? Renegotiating the

space of philosophical explanation". *Journal of Philosophy* 93: 483–515.

McGinn, Colin. (1991). *The Problem of Consciousness*. Oxford: Blackwell Publishing.

McGinn, Colin. (2004). "Inverted First-person Authority". *The Monist* 87 (2): 237-254.

McKinsey, Michael. (1991). "Anti-Individualism and Privileged Access". *Analysis* 51: 9-16.

McKinsey, Michael. (2002). "Form of Externalism and Privileged Access". *Philosophical Perspectives* (16): 199-224.

McTaggart, J. Ellis. (1908). "The Unreality of Time". *Mind* 17, 68: 457-474.

Mellor, D.H. (1981). *Real Time*. Cambridge: Cambridge University Press.

Mellor, D.H. (1991). "I and now" in Mellor, D.H., *Matters of Metaphysics*, Cambridge: Cambridge University Press.

Mellor, D. H. (1998). *Real Time II.* London: Routledge.

Merlo, Giovanni. (2013). "Specialness and Egalitarianism". *Thought: A Journal of Philosophy* 2 (3): 248-257.

Merricks, Trenton. (2006). "Goodbye Growing Block" in in Zimmerman, Dean (Ed.), *Oxford Studies in Metaphysics (Vol. 2)*, Oxford: Clarendon Press.

Mill, John Stuart. (1872). *An Examination of Sir William Hamilton's Philosophy* (fourth edition), London: Longman, Green, Reader, and Dyer.

Moore, Adrian W. (1997). *Points of View.* Oxford: Clarendon Press.

Morgan, Daniel. (2012). "Immunity To Error Through Misidentification: What Does It Tell

Us About The De Se?" in Prosser S. and F. Recanati (Eds.), *Immunity to Error Through Misidentification: New Essays*, Cambridge: Cambridge University Press, 2012: 104-123.

Nagel, Thomas. (1974). "What is it like to be a bat?" *Philosophical Review*, 83: 435-456.

Nagel, Thomas. (1979). "Panpsychism" in Nagel, T., *Mortal Questions*. Cambridge: Cambridge University Press.

Nagel, Thomas. (1986). *The View From Nowhere*. Oxford: Oxford University Press.

Nagel, Thomas. (1998). "Conceiving the Impossible and the Mind-Body Problem". *Philosophy* 73: 337-352.

Nemirow, Laurence. (1990). "Physicalism and the Cognitive Role of Acquaintance" in William Lycan (Ed.), *Mind and Cognition: A Reader*, Oxford: Blackwell (pp. 490-499).

Newman, Lex. (2010). "Descartes' Epistemology", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2010/entries/descartes-epistemology/>.

Nichols, Shaun and Stephen P. Stitch. (2003). *Mindreading*. New York: Oxford University Press.

Nichols, Shaun and Stephen P. Stitch. (2004). "Reading One's Own Mind: Self-Awareness and Developmental Psychology" in Stanton, R., Ezcurdia, M. and C. Viger (Eds.), *New Essays in Philosophy of Language and Mind, Canadian Journal of Philosophy,* vol. 30, Calgary: University of Calgary Press.

Nida-Rümelin, Martine. (1998). "On Belief About Experiences: An Epistemological Distinction Applied to the Knowledge Argument". *Philosophy and Phenomenological*

Research 58 (1): 51–73.

Nida-Rümelin, Martine. (2011). "The Conceptual Origin of Mind-Body Dualism" in Coliva, A. (Ed.), *Self and Self-Knowledge*, Oxford: Oxford University Press, 2011

Ninan, Dilip. (2009). "Persistence and the First-person Perspective". *Philosophical Review* 118 (4): 425-464.

O'Callaghan, Casey. (2008). "Seeing What You Hear: Cross-modal Illusions and Perception". *Philosophical Issues* 18 (1): 316-338.

O'Leary-Hawthorne, John and Andrew Cortens. (1995). "Towards ontological nihilism". *Philosophical Studies* 79: 143-165.

Papineau, David. (2001). "The Rise of Physicalism" in Gillett, C. and B. Loewer, *Physicalism and Its Discontents*, Cambridge: Cambridge University Press (pp. 3-36).

Papineau, David. (2002). *Thinking about Consciousness.* Oxford: Oxford University Press.

Parfit, Derek. (1998). "Divided Persons and the Nature of Persons" in Van Inwagen, P. and D. W. Zimmerman (Eds), *Metaphysics: The Big Questions*, Oxford: Blackwell (pp. 310-316).

Parsons, Josh. (2002). "A-Theory For B-Theorists". *The Philosophical Quarterly* 52, 206: 1-20.

Peacocke, Christopher. (1999). *Being Known*. Oxford: Oxford University Press.

Perry, John. (2001). "Time, consciousness and the knowledge argument" in Oaklander, L.N. (Ed.), *The Importance of Time*, Dordrecht: Kluwer Academic Publishers.

Perry, John. (2006). "Mary and Max and Jack and Ned" in Zimmerman, D. (Ed.), *Oxford*

Studies in Metaphysics 2006, vol. 2 (pp. 80-89).

Prior, Arthur N. (1959). "Thank Goodness That's Over". *Philosophy* 34, 128: 12-17.

Prior, Arthur N. (1962). "Changes in Events and Changes in Things". *The Lindley Lecture*. Lawrence: University of Kansas Department of Philosophy.

Prior, Arthur N. (1968). *Papers on Time and Tense*. Oxford: Clarendon Press.

Prior, Arthur N. (1968a). "Now". *Noûs* 2, 2: 101-119.

Prior, Arthur N. (1968b). "Egocentric Logic". *Noûs* 2, 3: 191-207.

Prior, Arthur N. (1968c). "Quasi-propositions and Quasi-individuals" in Prior (1968).

Prior, Arthur N. (1968d). "The Logic of Ending Time" in Prior (1968).

Prior, Arthur N. (1970). "The Notion of the Present," *Studium Generale* 23: 245-248.

Prior, Arthur N. (1998). "Some Free Thinking About Time" in Van Inwagen, P. and D. W. Zimmerman (Eds), *Metaphysics: The Big Questions*, Oxford: Blackwell (pp. 104-107).

Prior, Arthur N. and Kit Fine (1977). *Worlds, Times and Selves*. Amherst: The University of Massachussets Press.

Pryor, James. (1999). "Immunity To Error Through Misidentification". *Philosophical Topics* 26 (1-2): 271-304.

Pryor, James. (2000). "The Skeptic and the Dogmatist". *Noûs* 34 (4): 517-549.

Putnam, Hilary. (1970). "Is Semantics Possible?". *Metaphilosophy*, 1.

Rayo, Agustin and Stephen Yablo. (2001). "Nominalism Through De-Nominalization".

Noûs 35 (1): 74-92.

Rea, Michael C. (2003). "Four-Dimensionalim" in Loux, M.J. and D.W. Zimmerman (Eds.), *The Oxford Handbook of Metaphysics*, Oxford: Oxford University Press.

Recanati, François. (2006). "Indexical Concepts and Compositionality" in Carpintero, G. and J. Macia (Eds), *Two-dimensional Semantics*, Oxford: Clarendon Press (pp. 249-257).

Recanati, François. (2007). *Perspectival Thought: A Plea for (Moderate) Relativism*. Oxford: Oxford University Press.

Recanati, François. (2012a). *Mental Files*. Oxford: Oxford University Press.

Recanati, François. (2012b) "Immunity to Error Through Misidentification: What It Is And Where It Comes From" in Prosser S. and F. Recanati (Eds.), *Immunity to Error Through Misidentification: New Essays*, Cambridge: Cambridge University Press, 2012: 180-201.

Remes, Pauliina. (2008). "Inwardness and Infinity of Selfhood: from Plotinus to Augustine" in in Remes, P. and J. Sihvola (Eds.), *Ancient Philosophy of the Self*, Springer (pp. 155-176).

Richard, Mark. (1981). "Temporalism and Eternalism". *Philosophical Studies* 39: 1-13.

Rosenkranz, Sven. (2012). "In Defence of Ockhamism". *Philosophia* 40: 617-631.

Russell, Bertrand. (1910). "Knowledge by Acquaintance and Knowledge by Description". *Proceedings of the Aristotelian Society*, 11: 108–128.

Russell, Bertrand. (1986). *The Philosophy of Logical Atomism and Other Essays: 1914–19*, edited by J. G. Slater. London: George Allen and Unwin.

Ryle, Gilbert. (1949). *The Concept of Mind*. New York: Barnes and Noble.

Salmon, Nathan. (2003). “Tense and intension” in Smith Q. and A. Jokic (Eds.), *Time, tense, and reference ,* Cambridge, MA: MIT Press, (pp.107-154).

Sanson, David and Ben Caplan. (2010). “The Way Things Were”. *Philosophy and Phenomenological Research* 81 (1): 24-39.

Sartre, Jean-Paul. (1991). *The Transcendence of the Ego. An Existentialist Theory of Consciousness*. Traslated and annotated by Williams, F. and R. Kirkpatrick. New York: Hill and Wang.

Schaffer, Jonathan. (2003). “Is there a fundamental level?”. *Noûs* 37, 3: 498-517.

Schaffer, Jonathan. (2008). “The Metaphysics of Causation”. *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (Ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/causation-metaphysics/>.

Schaffer, Jonathan. (2009). “On what grounds what” in Chalmers, D., Manley, D. Manley and R. Wasserman (Eds.), *Metametaphysics*, Oxford: Oxford University Press.

Schwitzgebel, Eric. (2012a). “Introspection”. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2012/entries/introspection/>.

Schwitzgebel, Eric. (2012b). “Introspection, What?” in Smithies, D. and D. Stoljar (Eds.), *Introspection and Consciousness*, New York: Oxford University Press (pp. 29-48).

Searle, John R. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.

Searle, John R. (1997). “Reductionism and the Irreducibility of Consciousness” in Block et

al. (1997).

Shoemaker, Sidney. (1968). “Self-Reference And Self-Awareness”. *The Journal of Philosophy* 65 (19): 555-567.

Shoemaker, Sidney. (1994). ”Self-knowledge and ‘Inner Sense’. Lecture I: The object perception model”. *Philosophy and Phenomenological Research* 54: 249–269.

Sider, Theodore. (2001). *Four-Dimensionalism. An Ontology of Persistence and Time*. Oxford: Clarendon Press, .

Sider, Theodore. (2011). *Writing the Book of the World*. Oxford: Oxford University Press.

Smart, J. J. C. (1998). “The Space-Time World: an Excerpt from Philosophy and Scientific Realism” in Van Inwagen, P. and D. W. Zimmerman (Eds), *Metaphysics: The Big Questions*, Oxford: Blackwell (pp. 94-100).

Smith, Quentin. (1993). *Language and Time*. New York: Oxford University Press.

Soames, Scott. (2011). “True At”. *Analysis*, 71: 124–133.

Stalnaker, Robert. (1984). *Inquiry*. Cambridge: Cambridge University Press.

Stalnaker, Robert. (2012). *Mere Possibilities: Metaphysical Foundations of Modal Semantics*. Princeton, NJ: Princeton University Press.

Stanley, Jason. (2011). *Know How*. New York: Oxford University Press.

Stanley, Jason and Timothy Williamson. (2001). “Knowing How”. *Journal of Philosophy*: 98: 411-44.

Steward, Helen. (1997). *The Ontology of Mind: Events, Processes and States*. Oxford:

Clarendon Press.

Stoljar, Daniel. (2009). "Physicalism". *The Stanford Encyclopedia of Philosophy (Fall 2009 Edition)*, Edward N. Zalta (Ed.), URL = <http://plato.stanford.edu/archives/fall2009/entries/physicalism/>.

Swinburne, Richard. (1986). *The Evolution of the Soul*. Oxford: Oxford University Press.

Thomason Richmond H. and Robert Stalnaker. (1973). “A Semantic Theory of Adverbs”. *Linguistic Inquiry* 4 (2): 195-220.

Thornton, Stephen P. (2004). “Solipsism And The Problem Of Other Minds”. *Internet Encyclopedia of Philosophy*, Bradley Dowden (Ed.), URL = <http://www.iep.utm.edu/solipsis/>.

Tye, Michael. (2000). *Consciousness Color and Content.* Cambridge, MA: MIT Press.

Tye, Michael. (2009). "Qualia". *The Stanford Encyclopedia of Philosophy (Summer 2009 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2009/entries/qualia/>.

Unger, Peter. (2004). “The Mental Problem of the Many” in Zimmerman, Dean (Ed.), *Oxford Studies in Metaphysics (Vol. 1)*, Oxford: Oxford University Press.

Van Gulick, Robert. "Consciousness". (2009). *The Stanford Encyclopedia of Philosophy (Winter 2009 Edition)*, Edward N. Zalta (Ed.), URL = <http://plato.stanford.edu/archives/win2009/entries/consciousness/>.

Williams, Bernard. (1973). *Problems of the Self: Philosophical Papers 1956-1972*. Cambridge: Cambridge University Press.

Williamson, Timothy. (1996). “Cognitive Homelessness”. *The Journal of Philosophy* 93 (11): 554-573.

Williamson, Timothy. (1998). “Bare Possibilia”. *Erkenntnis* 48: 257-273.

Williamson, Timothy. (2000). *Knowledge and Its Limits*. Oxford: Oxford University Press.

Williamson, Timothy. (2002). “Necessary Existents” in A. O’Hear (Ed.), *Logic, Thought and Language*, Cambridge: Cambridge University Press: 233-251.

Williamson, Timothy. (2013). *Modal Logic As Metaphysics*. Oxford: Oxford University Press.

Wittgenstein, Ludwig. (1958). *The Blue and Brown Books*. Oxford: Blackwell.

Wittgenstein, Ludwig. (1961). *Tractatus Logico-Philosophicus*. Translated by Pears, D.F. And B.F. McGuinness. London: Routledge and Kegan Paul.

Wittgenstein, Ludwig. (1986). *Philosophical Investigations*. Translated by Anscombe, G.E.M. Oxford: Blackwell.

Woolf, Raphael. (2008). “Socratic Authority” in Remes, P. and J. Sihvola (Eds.), *Ancient Philosophy of the Self*, Springer (pp. 77-108).

Wright, Crispin. (1998). “Self-Knowledge: The Wittgensteinian Legacy” in Wright C, B. Smith and C. McDonald (Eds.), *Knowing Our Own Minds*, New York: Oxford University Press, 1998: 13-45.

Wright, Crispin. (2002). “(Anti)-sceptics, Simple and Subtle: Moore and McDowell”. *Philosophy and Phenomenological Research* 65: 330–48.

Wright, Crispin. (2012). “Reflections on François Recanati’s ‘Immunity To Error Through

Misidentification:What It Is And Where It Comes From'" in Prosser S. and F. Recanati (Eds.), *Immunity to Error Through Misidentification: New Essays*, Cambridge: Cambridge University Press, 2012: 247-280

Yablo, Stephen. (2002). "Coulda, Woulda, Shoulda" in Gendler, T.S. and J. Hawthorne (Eds.), *Conceivability and Possibility*, Oxford: Oxford University Press (pp. 441-492).

Zimmerman, Dean. (2005). "The A-Theory of Time, the B-Theory of Time and 'Taking Tense Seriously'". *Dialectica* 59, 4: 401-457.