



A Review of Brie Gertler (ed.) *Privileged Access: Philosophical Accounts of Self-Knowledge*

Philippe Vellozzo
EHES
Institut Jean Nicod
UMR 8129
1 bis avenue de Lowendal
F-75007 Paris
philippe_vellozzo@club-internet.fr

PSYCHE 11 (6), September 2005

REVIEW OF: *Privileged Access: Philosophical Accounts of Self-Knowledge*, B. Gertler (ed.), 2003, Aldershot: Ashgate Publishing.

Privileged Access provides a very valuable survey of contemporary philosophical views and issues on the classical topic of self-knowledge (knowledge of one's own mind). The book is in part an anthology of previously published papers, but it also contains nine new essays, most of which deal directly with the issue of privileged access to one's own mental states.

What is privileged access? There is a striking intuitive contrast between self-knowledge, on the one hand, and knowledge of other minds and of the external world, on the other hand. Our epistemic access to the external world, for instance, is perceptual and inferential. I can see directly that the cat is now in the kitchen or I can see footprints leading to the kitchen and thereby infer that the cat is now in the kitchen. We are highly fallible about the external world. But we seem to be more or less *infallible* about (at least some of) our own mental states. It is not easy to imagine that one has a false belief about one's current sensations, for example. If one believes that one feels cold, then it seems to follow that one feels cold. Moreover, our epistemic access to the mental states of other people is often taken to be indirect: One has to infer, it is said, what others think and feel from their behavior. But one clearly needs no such inference in order to know what one thinks and feels and in this sense, one knows one's own mind *directly*. Even if the latter difference is exaggerated, intuitively it is clear that one is normally in a *better position* than other people to know what one thinks and feels. So we seem to enjoy a rather special or "privileged" epistemic access to our own mind, when we compare it with the access we enjoy to the external world or to the other minds. (See the editor's introduction for a more complete characterization and a short discussion of these different ideas.)

It cannot be doubted seriously, I think, and all contributors in the present volume agree, that we enjoy, at least, *some kind of* “privileged” access to *some of* our own mental states. The problem of privileged access, then, is twofold. There is first a descriptive problem. The common-sense view of self-knowledge says that one has privileged access to one’s own mental states. But how are we to understand this claim? There is no general agreement concerning its *extension* and its precise *meaning*. (One may even question the unity of the phenomenon at issue, for it is not clear that we enjoy the same kind of access to all the mental states to which we seem to have privileged access.) Hence, the following questions require clear and well-developed answers:

Do we enjoy privileged access only to mental states? Do we enjoy privileged access to all of our mental states? If we have privileged access only to some of our mental states, to which mental states do we enjoy such access?

In what sense and to what extent is one’s epistemic access to—some of—one’s mental states “special” or “privileged”? Which of the intuitive differences between self-knowledge and other types of knowledge mentioned in the preceding paragraph are illusory and which are real? Is privileged access a simple matter of fact or is it, somehow, necessitated by the nature of the mental?

There is also an explanatory problem. If it is a fact that one enjoys privileged access to one’s own mind, then (dualism apart) it is certainly not a brute fact. If the intuitive differences between self-knowledge and other types of knowledge are not illusory, they cry out for explanation. One needs to explain why self-knowledge, as common sense pictures it, is so unlike ordinary empirical knowledge. The problem is especially acute for those who seek a broadly naturalistic account of the mental, for there is a *prima facie* tension between the scientific approach and privileged access. It is unclear that one would expect human beings to enjoy such privileged access from a scientific, e.g. evolutionary, point of view. Hence, some philosophers seek to reconcile privileged access with a naturalistically acceptable account of self-knowledge, while others, after Descartes and his followers, argue from the truth of (a strong version of) the claim of privileged access to the falsity of materialism. (See Gertler’s contribution for an interesting attempt to revive his argument.)

The problem of privileged access arises for all kinds of mental states (excluding unconscious states). It is a very general problem, but I will, for the most part, restrict my discussion to what I take to be an important aspect of it, namely the issue of privileged access to one’s own conscious *experiences* (mental events that have a qualitative or “phenomenal” character). Since it is widely agreed that we enjoy privileged access to our own experiences, theories of conscious experience that are incompatible with privileged access to one’s own experiences are in deep trouble. Hence, privileged access puts severe constraints on philosophical theorizing about “phenomenal” consciousness. Several new essays deal explicitly with this more specific issue, focusing on a particular view of experience known as *representationalism*. I will mainly comment on these essays in the following. I begin with a few general points about knowledge of one’s own experiences.

1. Knowing One’s Own Experiences

Common sense takes self-knowledge in general to be a *cognitive achievement*, that is, a type of knowledge on a par with our ordinary knowledge of the environment. I believe

that this common-sense view is particularly plausible regarding knowledge of one's own experiences. So I will put to one side, without argument, the view of those who believe that self-knowledge is not a cognitive achievement and the view of those who claim that self-knowledge is so unlike ordinary empirical knowledge that it should not be explained on the same model, be it observational or inferential (see the contributions of Bar-On & Long, Moran, Shoemaker, and Wright, for very different ways to develop these ideas). There are lots of problems with these two views (they are very counter-intuitive, they seem to threaten full-blown realism about the mental, etc.), and I believe that these problems are more acute when it comes to knowledge of one's own experiences. Imagine that you feel an itch in your right foot, for example. You are in a position to know quite precisely (even if you cannot describe it to someone else) what you feel in your right foot, *as long as you attend to it*. Things would be different if you were momentarily distracted, if you had, say, a sudden and violent pain in your left leg. It seems that you are, in this kind of case, "in cognitive touch" with something, for your attention is involved, as it is in your ordinary knowledge of the environment. Hence, one can have false beliefs about one's experiences because one's attention has been distracted or disrupted. Moreover, we can voluntarily attend to our own experiences (to their "phenomenal character"). As Boghossian (this volume, p. 76) puts it, "one can decide how much attention to direct to one's thoughts or images, just as one can decide how much attention to pay to objects in one's visual field." (See also Lycan, this volume, pp. 21-22) This issue of the involvement of attention in our first-person access to experiences is important (see the short discussion of the "transparency" of experience in section 2). I don't think that it makes sense to say we can attend to our own experiences in a way similar to the way in which we attend to external objects, as the quotation from Boghossian suggests, but we might avoid this problem by distinguishing between the *object* and *occupant* of attention (Peacocke, this volume, pp. 83-89).

For now, I will take knowledge of one's own experiences to be a standard, even if quite special, cognitive achievement, which means roughly that it is either observational (perceptual? quasi-perceptual?) or inferential.

At first sight, our introspective (or first-person) access to conscious experiences is very different from our introspective access to other types of mental states. I take it to be evident that we don't enjoy privileged access to unconscious states. But experiences are not the only type of mental state to which we claim to have privileged access. There are also so-called "propositional attitudes", prominently beliefs and desires. What distinguishes our epistemic access to the latter from our epistemic access to experiences? The answer, in a nutshell, is that (1) experiences are *occurrences* ("events", "episodes") but beliefs, desires and other propositional attitudes are not (necessarily) occurrences, and that (2) this determines the kind of epistemic access one may enjoy to them. A visual experience of a blue sky necessarily has a concrete impact on one's "mental life". But you can truly be said to believe, e.g., that the earth is a planet, or that you like chocolate mousse, while you are thinking about nothing at all or simply while you are asleep. Propositional attitudes are defined in *dispositional* terms, that is, roughly, in terms of their typical causes and effects. They need not be manifested. So, in itself, a belief (or a desire) is not a mental occurrence (while the internal manifestation of a belief, what is sometimes called a "judgment", surely is). Consequently, propositional attitudes do not have phenomenal character, for it is something like a conceptual truth that dispositions

don't have phenomenal character. Experiences, on the other hand, have a phenomenal character (a qualitative dimension) "by definition": there is "something it is like" for the subject to feel pleasure and pain, to have perceptual (e.g. visual) experiences, to feel sad or ashamed.

Actually judging that *p* is plausibly not necessary for being in a position to know that one believes that *p*. I may know that I believe that I have no car, without having my belief manifested in a judgment, e.g. I don't have to ask myself whether I believe that I have no car and answer that question (that is, judge that I have no car) in order to know that I believe that I have no car. But, since experiences are necessarily occurrences, first-person access to them necessarily is epistemic access to an occurrence and differs from first-person access to propositional attitudes, for it entails that something is now happening in one's own mind and that there is now a change in "what it's like" to be one. This explains, incidentally, why attention is so much involved in introspection of experiences. (Note that nothing that has been said so far suggests that our introspective access to experiences is better, in one way or other, than our introspective access to propositional attitudes.)

The claim that we enjoy (some kind of) privileged access to our own experiences is widely accepted. Intuitively, this is straightforward. It is plausible to hold that, in normal conditions, a rational subject does not have false introspective beliefs about what is consciously happening in her own mind. For instance, if *S* is a normal rational subject and *S* believes sincerely "from within" that she now feels pain in her back, it seems to follow that *S* feels pain in her back. If *S* is not absolutely infallible about the phenomenal character of her own experiences, then *S* is at least in a better position than anyone else to know "what it is like" for her to have such and such experience. (It may be impossible to *tell* others precisely what one feels, but that is clearly a different issue.) I should add that a false introspective belief about phenomenal character seems possible only insofar as the subject herself is in principle able to realize (after reflection) that her introspective belief is false. The very notion of phenomenal character seems to entail this much "first-person authority".

It is also difficult, again intuitively, to see how a normal rational subject could err in believing that she has now some experiences, that she *sees* *x*, *hears* *y* or *feels* *z* (and that she does not see *z*, hear *x* or feel *y*). Once more, even if this kind of introspective error is possible, one is nevertheless better placed than anyone else to know whether one has an experience and the kind of experience one has (e.g. its perceptual modality, if the experience is perceptual). Moreover, conscious occurrence in a normal rational subject seems to be almost a sufficient (if not necessary) condition for first-person knowledge that one has an experience of such and such a type. This intuition is sometimes expressed as the idea that conscious states are "self-intimating": If one has a conscious state *C* (and is rational), then one believes that one has *C*. So not only does one seem to enjoy privileged access to the precise phenomenal character of one's experiences (a point made, e.g. by Dretske and by Sosa, this volume), but one also seems to enjoy privileged access to the fact that one has now an experience of a certain type (visual, as opposed to tactile, say).

These intuitions might turn out to be deep cognitive illusions about ourselves, but a theorist who disagrees with them cannot refuse to take them seriously, for opposing

these intuitions will be counted as a high cost of her theory. One theory that apparently threatens privileged access to one's own experiences is the *externalist representationalism* eloquently advocated by Fred Dretske, William Lycan and Michael Tye. The compatibility of externalist representationalism with privileged access is then one important focus of the book. Five essays deal directly with this issue, to which I now turn.

2 Externalist Representationalism and Introspection

The notion of phenomenal character is sometimes taken as a primitive (irreducible), but many philosophers seek to explain phenomenal character in naturalistic, non phenomenal terms. One influential and promising reductive view of phenomenal character is *representationalism*. Representationalism—which I take here, for terminological simplicity, to be equivalent to what Tye calls “strong representationalism”—“aims to tell us what phenomenal character *is*” (Tye, this volume, p. 31), and it tells us that phenomenal character just is representational content of a certain sort. The representational content of an experience is what that experience represents. Visual experiences, for example, normally represent objects and properties (colors, shapes, distances, etc.) that are part of the environment. What it's like visually for subject S to see red strips on this white wall is exhausted by what her current visual experience represents in its own distinctive way (i.e. red strips on this white wall seen from a certain perspective, and so on).

Representationalism is usually (but not necessarily) held in conjunction with *phenomenal externalism*: the thesis that the representational content of experiences, hence phenomenal character, is externally determined, i.e. that it depends on present or (more plausibly) past causal relations between a creature and its environment. This conjunction is externalist representationalism. I will call it “ERQ” (“Externalist Representationalism about Qualia”), after Levine (this volume). ERQ entails that two subjects of experience S1 and S2 may be in exactly the same neural states (they may be neurobiological, and even molecule-to-molecule, duplicates) and have yet different experiences, for they are in different environments or have different histories. This is counter-intuitive, but maybe only so. (There are well-known other problems with ERQ, but here I will only be concerned with those related to privileged access.)

What do proponents of ERQ have to say about one's access to one's own experiences? ERQ is in principle compatible with a variety of accounts of our first-person access to experiences. But the “transparency” of experience is admittedly one very strong motivation for the view that phenomenal character is representational content, and the appeal to transparency can be turned into an argument for this view (see Tye, this volume, for such an argument). The intuition that experiences are “transparent” is that, if one introspects one's current visual experience, for instance, one “sees”, as it were, “right through it”. One simply “sees” the properties exemplified by the objects that one's experience is of. Introspective access to an experience involves attention, as has already been said, but this attention is directed outwards, to the external world so to say, even if the experience is illusory. (More precisely, it is directed to that part of the external world that the experience represents in its own way.) Since transparency is the main intuitive

support for externalist representationalism, I believe defenders of ERQ should offer a theory of introspection that gives prominence to it.

Nevertheless, some advocates of ERQ (e.g. Lycan, 1996) favor an “inner sense” view of introspection, according to which introspection is a quasi-perceptual faculty, involving attentional processes directed to experiences as such, that is, *qua* inner happenings. Introspection is then said to *justify* one’s beliefs about one’s own experiences in a way similar to the way perception (or perceptual attention) justifies one’s beliefs about one’s external environment. Whatever the intrinsic merits of this approach, the problem is, I think, pretty clear: If experience is transparent, then it seems natural to deny that introspection is a quasi-perceptual faculty directed inwards. So the “inner sense” view is entirely at odds with the spirit of ERQ, which leads quite naturally to the opposite view.

Moreover, the “inner sense” view has problems of its own. The common-sense idea of privileged access plausibly requires more than the mere natural fact that we are equipped with “internal scanners”. There seems to be a tight conceptual (hence not simply causal) link between self-knowledge and our distinctively human type of rationality. This counts against any view on which one has to observe or perceive one’s own mental states “from within”, in the way intended by defenders of the “inner sense” view, for perception is a causal process and, as such, it can be disrupted. It is then conceptually (and maybe even empirically) possible that, for whatever reason, a rational subject always misperceives her environment. If introspection were inner observation (or perception), as the term suggests, it would then be possible also that a rational subject’s beliefs about her mental states are systematically false. But, Shoemaker (this volume) argues, *this* is not conceptually possible. I think he’s basically right, even if his arguments for the conceptual impossibility of “self blindness” (the scenario just described) are not conclusive as they stand. “Self blindness” really does not *seem* conceptually possible. Moreover, Siewert (this volume) argues convincingly that the claim that “self blindness” is conceptually impossible can be defended on quite different grounds. Shoemaker’s positive view is that it is among the normal or typical effects of a (conscious) mental state *M* of a subject *S* that it gives rise, in *S*, to the true first-person belief that she is in *M*. Hence, in any humanly rational creature, second-order beliefs are logically *supervenient* on first-order ones: they are “nothing and above” first-order ones, “plus a certain degree of rationality, intelligence and conceptual capacity” (Shoemaker, this volume, p. 117). A view along those lines is very attractive, if incomplete. It can provide one with plausible constraints on acceptable theories of self-knowledge, but it seems insufficient as an account of self-knowledge, and particularly incomplete for knowledge of one’s own experiences (see Siewert, this volume, for some other worries). Nevertheless, the argument for the claim that introspection cannot be a purely empirical, perceptual or quasi-perceptual faculty seems to me to be sound.

Most advocates of ERQ endorse a different model of introspection, the “Displaced Perception Model” (DPM). This model is also empirical, but it can, according to Dretske (1995), validate a quasi-a priori type of infallibility, at least for representational content. On DPM, you come to know *that* your experience has a certain phenomenal character (= a certain content) *by* being aware *of* what the experience represents, of its externally individuated content (Dretske, 1995, Seager, 1999, Tye, this

volume, pp. 37-39). This “counter-etymological account of introspection” (Lycan, this volume, p. 15) has, I believe, many virtues. It is simple and parsimonious, and it is not entirely counter-intuitive, after reflection. So DPM seems promising, at least for someone who does not feel attracted by an “inner sense” view.

The problem is that DPM seems to entail that one has to *infer* one’s knowledge of one’s own experiences (e.g. that one sees that this flower is red) from one’s perceptual knowledge of the environment (that this flower is red). But, first, as Aydede points out, such an inferential model of introspection is *prima facie* implausible. It runs counter to the common-sense idea that one normally has direct (i.e. non inferential) knowledge of one’s own experiences. (Here, it could be replied that the kind of inference that is intuitively ruled out by the notion of privileged access is only inference from one’s behavior.) Moreover, Aydede argues convincingly, in my view, that one needs an appropriate “connecting” belief in order to infer that one is having *an experience* with a given content from that content itself, and that there is no plausible candidate, no such (justified) connecting belief, to be found. So this inferential model cannot be true.

I fully accept Aydede’s conclusion about the inferential DPM. But it could be questioned whether the Displaced Perception Model of introspection has to be inferential. Aydede seems to think so. He seems to think that any attempt to develop a non-inferential DPM is bound to fail, for a non-inferential DPM is either empty or incoherent. But I could find no real argument for this claim in his paper. I agree, nevertheless, that neither Dretske nor Tye has provided a detailed alternative construal of DPM. I confess I don’t know if such an alternative construal is possible, even in principle. Tye, for instance, is very explicit that his version of DPM is not inferential at all. He says: “If I am aware of certain external qualities, I do not need a background belief to be aware that I am undergoing an experience with a certain phenomenal character, once I introspect” (Tye, this volume, p. 38). He appeals instead to the idea of a simple “reliable process”. But this proposal is clearly *ad hoc*. Hence, DPM (and therefore ERQ?) is in trouble. Now, externalist representationalism will really be in trouble if it cannot deliver a plausible account of knowledge of one’s own experiences, an account, that is, that does not put into jeopardy the very idea of privileged access.

3 The Compatibility of ERQ with Privileged Access

But it seems that externalist representationalism is already in trouble. For in his stimulating and provocative contribution, Fred Dretske, a chief proponent of this view, suggests that, while we have privileged access to the content of our own conscious states (he goes as far as claiming that we are absolutely infallible about it), we don’t enjoy privileged access to the very fact that we have conscious states! Dretske is well aware that this may simply sound silly to some (myself included). We might, admittedly, learn to live with the fact (if it is one) that we don’t know in any direct or privileged way the “attitudinal aspect”, as he calls it, of our thoughts, that is, whether one’s thought that *p* is a belief or something else. But it is very difficult to see how we could lack direct or privileged access to our being conscious. Yet Dretske argues that nothing he can think of might explain *how* we know introspectively that we are conscious, and then claims that this gives us reasons to doubt that we really know introspectively (hence, *in any privileged way*) that we are conscious. He claims that the problem of accounting for our

knowledge of our own experiences is the same whether or not one endorses externalist representationalism. But this is far from clear, as we will see.

Dretske's view is clearly in conflict with the common-sense idea of privileged access to one's own conscious states. But one has to agree with him, I think, that "there is nothing we perceive that tells us we are conscious" (p.3). He shows also quite convincingly that the fact that pains and the like are "states we are necessarily aware of" does not tell us we are conscious either. He considers, finally, the idea that we know we have experiences by a direct, quasi-perceptual, awareness of them (the "inner sense" view) and finds it wanting. "I did not become aware of the fact that I have conscious experiences by an awareness of the conscious experiences themselves in the way I become aware of the fact that there is beer in the fridge by seeing the beer." (p.8). But, fortunately, even if we cannot give, for the moment, a complete and plausible answer to the question that serves as the title of his contribution ("How do you know you are not a zombie?"), we still have reasons to resist the conclusion that we enjoy no privileged access to the fact—and don't even know by introspection—that we are conscious.

Dretske aptly distinguishes between two kinds of difference one might be interested in: the kind of difference that exists between **seeing x** and **seeing y**, on the one hand, and the kind of difference that exists between **seeing x** and **not seeing x**, on the other hand. But there is a further important kind of difference that is not mentioned by Dretske. This is the kind of difference that exists between **seeing x** and **visualizing x** (that is, imagining x visually). Now, the latter difference is analogous to that between **believing that p** and, say, simply **wondering whether p**. It is then the perceptual equivalent of a difference in "attitudinal aspect". Now, one might grant that we have no real privileged access to the "attitudinal aspect" of our conscious states. One is sometimes simply unable to tell whether one is seeing or visualizing x. The conditions in which this might happen need not even be very unusual (e.g. daydreaming). But it is more difficult to believe that one is not introspectively aware of the perceptual modality of one's own perceptual experiences. And it is still more difficult to believe that we have no introspective access to whether or not our own mental states are conscious. My point is that it does not follow from the alleged fact that we enjoy no real privileged access to the "attitudinal aspect" of our conscious states that we do not enjoy privileged access to these states being conscious. One may not be in a position to know whether one sees x or visualizes x and yet be in a position to know that one has a *conscious* state (even maybe a "visual", or at least, perceptual, conscious state—an experience involving colors for example).

This may suggest the reply that consciousness is something like the "Ur-attitude", the more general experiential attitude. I really doubt that this is true. Or, if it is true, then it is trivially true. The trivial claim would be that consciousness is the property exemplified in all conscious episodes. But this, while true, does nothing to substantiate the claim that the concept of consciousness is comparable to the concepts of belief and desire, except that it is applied to the phenomenal domain and that it is much more general. This is not a trivial claim, and it is arguably false. The very notion of "experiential attitude" seems to me of dubious coherence. Consciousness, arguably, is not an attitude, nor an "attitudinal aspect", at all, nor is it easily comparable to such. At the very least, an argument is needed for the claim that it is. But Dretske offers no argument

for that claim. So, if his reasoning rests on the assumption that consciousness is attitude-like, as it seems to, then his conclusion simply does not follow.

I agree with Dretske that experiences themselves are “invisible, to the introspective scanner” (p. 9). This is the point of the transparency intuition. But, assuming consciousness is not attitude-like, the possibility is then open that consciousness is somehow (implicitly) present in the content of any conscious experience. The intuitive appeal of DPM is that, for a rational subject S, everything required for knowing that she has experience e_1 is there *in the content of e_1* . It could be argued that, if reference is allowed to consciousness (or to e_1 itself) in the very content of e_1 , then the difficulty of finding an appropriate “connecting belief” for the inference from the content of e_1 to the fact that S has now e_1 vanishes. Externalist representationalism rules out this theoretical possibility (which I only suggest here), as it rules out other possibilities that might help answer Dretske’s question. If you hold that consciousness is a property that a state with the right kind of content might *fail* to possess, and hence that having the right kind of content is not sufficient for a state being conscious, then you will probably have quite a different idea of what is involved in our first-person access to consciousness. This is why I think that the claim that we enjoy no privileged access to the fact that we are conscious, given its initial implausibility, is unwarranted, and that ERQ is in trouble.

This is not the only problem with ERQ. It is *prima facie* difficult to see how something which is *externally* determined could be known directly “from within”, for its identity (hence, the truth of our beliefs about it) depends on facts about the environment of which the subject may have no idea. Schematically, we have three (apparently) conflicting propositions:

- (i) One has privileged access to the phenomenal character of one’s own experiences.
- (ii) Phenomenal character is externally determined representational content.
- (iii) The environment one happens to be in is not something to which one has introspective (hence privileged) access.

Note that this is *not* simply the problem, stated as a “principle” in Boghossian (this volume, p. 73), that “you cannot tell by mere inspection of an object that it has a given *relational* or *extrinsic* property.” Can this intuitive difficulty be turned into a conclusive argument against ERQ? In his interesting contribution, Joseph Levine argues that ERQ (the conjunction of representationalism and phenomenal externalism) is incompatible with a plausible condition on self-knowledge. Levine’s argument against ERQ is original in that it does not depend on the (controversial) claim that introspection delivers knowledge only of intrinsic, that is, non-relational, properties. He claims that there is a “crucial difference” between ERQ and standard externalism, as far as compatibility with self-knowledge is concerned, and that it is not externalism *per se*, but the “reduction of phenomenal character to content” that threatens self-knowledge (this volume, p. 48).

Levine makes use of the notion of a “Frege case”. Roughly, a “Frege case”, in the domain of thought, is a case in which a rational subject S has distinct A-thoughts and B-thoughts while, in fact, A *is* B. More precisely, it is a case in which:

S judges that A is F.

S judges that B is not F.

A is B.

The possibility of this kind of case is compatible with S's being rational since S may not know that A is B. For example, Oedipus judges that he's marrying Jocasta, but he believes also (and may judge) that he is not marrying his mother. He is not irrational, since he doesn't know that Jocasta is his mother. The orthodox conclusion, which Levine endorses, is that, in some sense, Oedipus's Jocasta-thought is not his mother-thought. Each involves, says Levine, a different thought-vehicle. The important point for the argument is the following. This difference of vehicle is something a rational subject normally knows "from within". I can know that two thoughts are of different types, says Levine, "because I can distinguish, merely from what's going on inside, that they involve different vehicles." (p. 49). So it is plausible to hold that, if a rational subject judges *introspectively* that two thoughts are different, her judgment cannot be false for reasons inaccessible to her.

The general moral Levine wants to draw from the existence of Frege cases is then that the truth of a subject's judgments about whether her thoughts (experiences) are identical or different cannot depend on inaccessible external facts, like the fact that Jocasta is Oedipus's mother. For, if these judgments depend on inaccessible external facts, then Frege cases will be ruled out. But remember that, on ERQ, phenomenal characters are externally determined. That means that, if experiences e1 and e2 represent the same external property, then they have the same phenomenal character, hence are of the same phenomenal type. More explicitly, ERQ entails that if R and Q are the same property and if e1 is a R-experience and e2 is a Q-experience, then e1 and e2 are the same experience, even if their subject judges introspectively that they are different.

Standard externalism says that the *content* of propositional attitudes depends on relations with the external (and, on some views, social) environment. It is in principle silent about *vehicles*, the identity of which may or may not be determined by external factors, depending on one's general view of vehicles and of their relation to contents. So, given a widespread internalism about vehicles, externalists may allow that some differences between mental states, namely differences of vehicles, are not dependent on external factors and therefore can be known merely by introspection. By contrast, ERQ is a thesis about the individuation of conscious experiences themselves and so, according to Levine, cannot allow that some differences between experiences do not depend on external factors. Hence, the only simple way to avoid the objection would be to deny that Frege cases are possible for judgments about our own experiences, but I agree with Levine that this reply would be ad hoc.

Levine maintains that, in the case of thought, "if one judges that one is entertaining two distinct thoughts, and nothing bizarre is going on inside, then one is entertaining distinct thoughts." (p. 50) Once generalized, this means that, in normal conditions, a rational subject is somehow *infallible* concerning differences between her own mental episodes (thanks to her introspective knowledge of vehicles?). So the argument appeals to a "plausible condition on self-knowledge" which is something like the following:

If S is a normal rational subject and S judges introspectively that mental episodes M and M* are different, then M and M* are different.

Now, a defender of ERQ may object that this condition is too strong and that Levine's argument begs the question against ERQ. The condition is not self-evident, one may argue, for it is not entirely clear that our access to the phenomenal character of our own experiences is as reliable as Levine takes it to be. (See, for instance, Lycan's attack on what he calls the "hyperinfallibility" of introspection, this volume, p. 20). So, an argument is clearly needed here and Levine offers no argument. Yet such an objection is not entirely convincing, I believe, for we really do *seem* to enjoy privileged access to differences in phenomenal character. And if one is a representationalist who believes, like Dretske, that we are infallible about the content of our own experiences, then why not grant that we are also infallible about phenomenal (= representational) differences? Moreover, one could concede that we're not absolutely infallible about phenomenal differences, but still insist that ignorance of remote external facts cannot plausibly explain a false introspective judgment that two experiences are different. Surely, *this* needs to be argued. But it does not seem to me very difficult to find an argument for this slightly weaker condition, since it is far from clear, intuitively, how ignorance of external facts could lead us into error concerning differences between our own conscious experiences. So while Levine's argument against ERQ is not entirely conclusive as it stands, I believe that a conclusive argument along the same lines could be given for the claim that externalist representationalism is incompatible with privileged access to phenomenal differences. Whether this is a serious problem for externalist representationalism will of course depend, in the end, on how much theoretical respect is due to our ordinary conception of privileged access.

Most philosophers rightly reject the Cartesian reading of the claim that one enjoys privileged access to one's own experiences. But we normally do have a rather special epistemic access to our own experiences, even if we're not absolutely infallible about them. Several essays in this volume deal explicitly with an influential theory of conscious experience, namely externalist representationalism, that conflicts with this plausible claim about self-knowledge. I have tried to show that this is a serious problem for externalist representationalism: it seems to be in the very nature of conscious experiences that subjects, if rational, enjoy privileged access to their own experiences. One may add that the scientific study of consciousness itself requires taking introspective reports at face value. Our review suggests, then, that even a weak reading of the claim of privileged access puts severe constraints on acceptable theories of consciousness.

References

- Dretske, F. (1995), *Naturalizing the Mind*, Cambridge, MA: MIT Press, Bradford Books.
- Lycan, W. (1996), *Consciousness and Experience*, Cambridge, MA: MIT Press, Bradford Books.
- Seager, W. (1999), *Theories of Consciousness: An Introduction and Assessment*, London: Routledge.