

## Self-knowledge: the Reality of Privileged Access

Crispin Wright

Paul Snowdon's [2012] develops a range of careful and interesting criticisms of ideas about the problem of self-knowledge, and about what I interpreted as the broad contribution to it made by Wittgenstein's later work, that I presented in Whitehead lectures at Harvard almost twenty years ago.<sup>1</sup> Snowdon questions whether Wittgenstein's characteristic focus upon the linguistic expressions of self-knowledge holds out any real prospect of philosophical progress, and charges that my discussion is guilty in any case of distortion and over-simplification of the 'data', whether conceived as linguistic or otherwise, that set the problem of self-knowledge in the first place. In this paper, I take the opportunity to respond.

§1. Although some philosophers have preferred to think otherwise,<sup>2</sup> it is not philosophy but part of the ordinary folk notion of the mental, enshrined in literature and drama, that each of us stands in a special relationship, denied to others, to our own mental lives—that a significant range of our mental states and attributes are directly available to us and only indirectly available to others, that “You cannot really know what another is thinking”, for instance, whereas of one’s own occurrent thoughts one cannot but be aware. This special relationship seems to embrace each of sensation, mood, emotion, belief, desire, fear, intention, action (what I am doing), memory (what I am remembering), perception (what I am seeing), thought (what I am thinking), imagination, (idiolectic) meaning ... the list goes on.

Since the furore provoked by Michael McKinsey's [1991], it has often, though not always, been accepted that this *privileged access* we seem to have to our mental states—specifically, our privileged access to a wide class of our *attitudinal* states—collides with the forms of externalism about mental content championed by Burge, Putnam and others,<sup>3</sup> whose

---

<sup>1</sup> The lectures, entitled “The Problem of Self-Knowledge”, were delivered at Harvard in the Spring of 1996. Since they were first published as chapters 10 and 11 of Wright [2001], I shall so refer to them throughout.

<sup>2</sup> See for example the introductory remarks in McDowell [1991].

<sup>3</sup> The *loci classici* are Putnam [1975] and Burge [1979].

acceptance has recently become something like orthodoxy. It is notable, though, that the case for collision—the "McKinsey paradox"—exploits only a very spare conception of privileged access: merely, that someone can know 'in the armchair' that they e.g. believe a proposition whose content is externally determined by matters like the nature of their physical environment or the history and practices of their speech community, yet know this by means involving no investigation of such matters. That we have privileged access to a wide class of our attitudinal states according to so spare a conception of privileged access may seem ungainsayable. So it is understandable that much of the literature has focused on the externalist component in the resulting aporia.<sup>4</sup> I myself am one of those who are unpersuaded that there is actually any collision in the first place.<sup>5</sup> However I do believe, as Snowden's discussion indeed brings out, that a more extended and attentive investigation of the other putative component—the nature and extent of privileged access—is long overdue.

As an initial approximation, and setting aside certain differences occasioned by the types of mental attribute concerned, privileged access seems to involve three separable features. First, much of the knowledge we possess of our own mental states is seemingly *immediate*: we do not in general, in coming to knowledge of our own psychological attributes, need to rely on the kind of evidence—what they say, how they act, how they look—on which we need to rely in coming to knowledge of others' or on which, *mutatis mutandis*, they have to rely in coming to know of ours. Immediacy here is at least *non-inferentiality*; whether it is something more is something that we will need to consider further below. However that may be, immediacy doesn't of itself imply the second feature: that our opinions about our own psychological attributes often carry a distinctive *authority*: that selves not only know differently to others, as follows from immediacy, but also generally know *better*. And the third feature — *salience*—reverses the direction of fit: selves tend to know what there is to know—selves' mental attributes do not, in general, elude their awareness.

These generalisations are, as I say, an initial approximation. Yet granting that they are even approximately true, and that privileged access, so outlined, is not chimerical, they set a philosophical challenge: to give an account of the self-other asymmetries that they characterise—to explain their source, or ground. Of course it is not transparent what a distinctively *philosophical* explanation should here accomplish. There is, however, a venerable—but deeply

---

<sup>4</sup> An excellent overview of the debate is provided in Ted Parent's article, "Externalism and Self Knowledge" in the *Stanford Encyclopaedia of Philosophy* at <http://plato.stanford.edu/entries/self-knowledge-externalism/>

<sup>5</sup> My most recent thoughts about this are in Wright [2011]. I won't go further into the issue here.

problematic—response to the *prima facie* asymmetries that springs to mind so naturally that it can seem constitutive of the subject matter. This is the idea, associated (perhaps erroneously) by much modern thought with Descartes, that each one's mental life constitutes a totally transparent *inner theatre*, with an audience, necessarily, of one. Others will thus need indirect evidence to suppose that something is happening on your inner stage, but you can just observe it. So of course you know best. And since there is total transparency, you will observe what is there to be observed.

The venerable idea is, of course, deeply problematic for several familiar reasons. To begin with, the appeal to a kind of inner observation involves some degree of distortion of the phenomenology of certain of the attributes concerned—for example, when you recognise that you believe something, there need be no distinctive state of consciousness involved, as there is with, say, a toothache or an itch. But more theoretically, the model plays a villain's part in generating the other minds problem. And its very coherence is put in serious question by the concerns about 'private language'—really, private conceptualisation—originated by Wittgenstein in the *Philosophical Investigations*.

The cardinal philosophical problem of self-knowledge, as conceived in the modern debates, has been to do better than the venerable response: to achieve a satisfactory perspective on what is correct about the idea of privileged access—what is special about a subject's self-impressions—which can head off the aporia of what (for convenience) we may call *Cartesianism*, while at the same time finding a place for (perhaps refined versions of) the seemingly undeniable asymmetries that motivate the venerable notion.

§2. The trio of *immediacy*, *authority* and *salience* allow of two quite different directions of interpretation. On the one hand, each of the three can, obviously, be viewed as a feature of the *cognitive relationship* in which a subject stands to her own mental attributes. On the other hand, and perhaps less obviously, they can be viewed as primarily *features of the discourse* in which we speak of our own and others' mental attributes. Following Ryle's example in *The Concept of Mind* I shall use the term 'avowal' to denote the kind of psychological self-ascription whose properties suggest privileged access. Since Ryle— and Wittgenstein—wrote, one tendency of analytical philosophical discussion of psychological self-knowledge has been to focus on its linguistic expression and on the first-third person asymmetries as reflected in characteristics of the competent use of avowals and the competent ascription of mental states to others. So conceived, as a phenomenon of psychological discourse, immediacy will have to do e.g. with the inappropriateness of requesting grounds for a subject's avowals, authority with the propriety of deference to what she has to avow, and salience will be located in the inappropriateness of professions of ignorance of one's own avowable states. And these points—no doubt needing

refinement to accommodate details of context and variations in the kinds of mental attribute concerned—are, when viewed from the standpoint of the 'linguistic turn', points not about epistemology but about the 'language-game', about our practices of ordinary psychological ascription.

So the would-be theorist of self-knowledge confronts a fork. What comes first here in the order of explanation: the linguistic practice, or the thoughts of the thinkers manifested in that practice? The problem of self-knowledge will look very different depending on how we choose. On the first option, we will tend to think of matters very much as recent writers such as Bar-On and Finkelstein<sup>6</sup> have continued to do, as centred upon explaining the distinctive marks of avowals. Our problem will be viewed as that of accounting for the “grammar” of avowals, in the respects in which it contrasts with that of the competent other-ascription of mental states; and a range of candidate explanations will then enter the fray, for instance those falling under the broad rubric of *Expressivism*, which are otherwise excluded. On the latter option, by contrast, the relevant features of avowals will, from the outset, be seen as reflections of aspects of the *epistemic character* of the self-directed thoughts they express. And it will seem overwhelmingly natural to suppose just what pre-philosophically we naturally do suppose: that selves characteristically know of the states that give rise to avowal in a way that involves no inference from independent reasons, and which is characteristically very secure; and that the states of the relevant kind are typically salient to their subjects. The problem will then be to account for these apparent epistemic advantages in a way that guards against a slide into Cartesian privacy and its associated nemeses.

§3. So much for stage setting. Snowdon’s principal critical contentions are, in summary, the following three:

First, he is sceptical about the value of the linguistic turn in the modern discussion sponsored by Ryle and Wittgenstein, holding that it stands unproven and presents “an as yet undecided but very important question.” Indeed, his discussion reads, more strongly, as arguing that the Wittgensteinian focus upon language is a misdirection, and that we should consider features of psychological self-knowledge, or features of beliefs about one's own psychological attributes, in their own right rather than concentrate on aspects of the discourse in which such knowledge and beliefs are apparently expressed.

Second, Snowdon makes complications for the taxonomy of avowals that informed my earlier treatment— the broad division between what I termed *phenomenal* avowals (typified by "I have a headache" and "There is a ringing in my ears") and *attitudinal* avowals (typified by "I

---

<sup>6</sup> Bar-On [2004] and Finkelstein [2003]

believe that my printer is ready" and "I fear he may not have survived".) I think there is justice in his criticisms here and will have more to say below about how better we might organise the 'data'.

Third, Snowdon argues that, when interpreted as on the linguistic side of the fork—i.e. as properties of competent practice in the language-game of self- and other- psychological ascription—immediacy, authority and salience do in any case not in fact hold good and involve confusions. Moreover, he continues, even after we correct for this and revert to more purely epistemic construals of the trio, they still emerge as incorrect and/or needing qualification.

I will elaborate on and assess these contentions in turn.

§4. The term “avowal”, employed with the sense, roughly, of: authoritative, non-inferential psychological self ascription, is used as if familiar in *The Concept of Mind*, and was evidently a term of art in debates in the philosophy of mind for some time before that work. Now, the problems posed by self-knowledge surely can be raised in a fashion that focuses on avowals without begging any important questions. We can ask, for example, why there are no correspondingly authoritative, non-inferential psychological *other* -ascriptions, and in posing that question, we leave open what the explanation might be—whether it might be centred on issues in epistemology and on the cognitive relationship that thinkers bear to their own mental states, or whether we might rather here have to do with an aspect of our linguistic practice that is, say, primitive and fundamental. Snowdon, however, is uneasy about the preoccupation with avowals. Why, he wonders, this focus on features of the way in which we *give expression to* self-knowledge, — why engage the problem in this way, rather than attend directly to “the phenomenon itself”? In the discussion of mine to which he is responding, it is true that the trio of properties of immediacy, authority and salience<sup>7</sup> are for the most part all interpreted as characteristics of speech acts—as pertaining to the conditions of propriety and impropriety of certain utterances and of reactions to them. Thus it is said to be inappropriate to question avowals in certain ways, absurd to profess certain kinds of doubt, required that other kinds of doubt assume a certain shape, and so on. Why, other than because in the grip of a certain post-Wittgensteinian dogma about philosophical method, should we engage the issues about self-knowledge by reflection on such normative claims about proper linguistic practice?

Well, I can only speak for myself. The explanation of the prominence of this type of mode of presentation of the data in my own work was less any commitment to a certain Wittgensteinian conception of philosophical method than a desire to bring out the choice point—the fork—noted above. If the explanandum is that selves' opinions about themselves have a

---

<sup>7</sup> There characterised as groundlessness, authority and transparency.

certain kind of characteristic immediacy and cognitive supremacy, the only admissible form of explanation is going to seem to be that selves do enjoy some kind of cognitive advantage, of which it then behoves us to give a model, or an account. But if the datum for explanation consists rather in the operation of a range of distinctive asymmetries and normative constraints on psychological discourse, then although an explanation in terms of cognitive advantage is *by no means thereby ruled out*, the possibility of other kinds of explanation—and indeed the ‘primitivist’ suggestion, which I take to be Wittgenstein’s, that there is in the end no explanation to give—becomes salient. Snowdon is quite right when he suggests that what I take to be Wittgenstein’s proposal—that the relevant features of avowals may be “grammatical”—fits naturally with a preoccupation with speech acts as a starting point, and is much less natural, or relevant-seeming otherwise. So far from skewing the discussion, however, I think that this way of setting it up makes for a much better understanding of Wittgenstein’s view than otherwise. That is not to admit that it should in any way predispose one towards Wittgenstein’s view. All the options remain on the table.

I do think it vital that, in attempting to evaluate Wittgenstein’s contribution to these questions, we keep in mind his later explicit conception of philosophical problems and philosophical method. His central idea, familiarly, is that philosophical puzzles are driven by misunderstandings of our discourse. In the present case, I take it that the key “misunderstanding” is, as suggested above, that certain superficial analogies between the ways that avowals are made and received and the way that reports of observation are made and received motivate a construction of our epistemic relation to the subject matter of the former on the model of a (variously idealized) version of our epistemic relation to the subject matter of the latter. This provides the shape of a diagnosis about where the Cartesian conception of the mind comes from. It represents it as an artefact of an anyway inappropriate form of theorizing: that of fashioning conceptions of the subject matter of a discourse, and of our relationship to that subject matter, in order to explain and underwrite its ‘grammar’. While it is of course an open issue whether this broad idea is fruitfully deployed in this case, or correct in general, it’s crucial to giving it at least a run for its money that one begin at the ‘level’ of language.

To set the record straight, then: I did not mean in the Whitehead lectures, to side with Wittgenstein’s view of these matters. But I did want to set things up in a way that allows his view to be heard. One can imagine a version of Snowdon’s complaint being levelled against the early pioneers of expressivism about ethical discourse: “Why the preoccupation with the way we *talk* morally, why don’t we just get on and study the nature of the moral facts and our cognitive relationship to them?” That response would simply render invisible the question that the ethical

expressivists wanted to raise, namely whether the “propositional surface”<sup>8</sup> of ethical discourse did not in fact mislead as to the nature of the role and content of that discourse, and whether indeed the whole notion of the “ethical facts” did not incorporate a metaphysical mistake. In the present case the possibility of mistake that is relevant would be not metaphysical but epistemological. But that possibility will simply not be in view if we start out locked into the assumption that the features of the discourse—the first-third person asymmetries—that strike us as surprising and in need of explanation are to be attributed to special cognitive advantages that selves enjoy in relation to their own mental states. I therefore reject Snowdon's suggestion that the preoccupation with avowals slants the discussion in Wittgenstein's favour. Rather it merely puts us in position to appreciate what, rightly or wrongly, he may have been driving at.

Before moving on, let me quickly address one objection Snowdon makes to the 'primitivist' view. Taken at the level of discourse, the three proposed marks of self knowledge are expressed in terms of inappropriateness, incongruity, absurdity, and so on. Thus, it is inappropriate to request that someone gives her grounds for a self-ascription of pain; incongruous to challenge such a self-ascription in normal circumstances; and absurd that somebody might profess themselves ignorant of whether they had a severe headache or not. Yet these are not, Snowdon correctly observes, the types of parameter of appraisal that are in general characteristic where primitive and constitutive rules of practice are involved. It would not, for example, be somehow *absurd*, or *incongruous*, if the Bishop were to be empowered with different ranges of movement in chess. It would just be a different game. In general, the suggestion would be, when a rule is primitive and groundless, then there will be, just for that reason, no space for any kind of basis or support for it—so that the idea of dropping or modifying it in certain respects should not clash with anything important and thus should involve no attendant absurdity or sense of impropriety. There is nothing to misrepresent or compromise.

Although this thought would sustain a more extended discussion that I shall give it here, I very much doubt that there is a telling objection to Wittgenstein's primitivism to be developed along these lines. The objection overlooks that the “rules of grammar” that are part of the Wittgensteinian road map are rules that fix our concepts of what is possible, that determine what *makes sense* to us. It is, for example, or so I would suggest, a good candidate for a Wittgensteinian “rule of grammar” that “knows”, at least as normally used, is factive. That factivity of knowledge is not something imposed on us from above, as it were. Factivity is something we built in to our notion of the highest cognitive accomplishment. No doubt there are

---

<sup>8</sup> Simon Blackburn's [1993] nice phrase.

reasons for that, but at the same time it is readily conceivable that there be communities whose notions of the ways a belief may be in good standing include none that actually requires factivity. Still, although the factivity of “knows” is in this sense a shallow fact, it's still the case that professions of knowledge of what is as a matter of fact false will strike us as absurd, rather than merely breaches of some groundless convention.

§5. In Wright [2001] I wrote that:

It seems safe to suppose that we must begin by distinguishing two broad classes of avowal. The first group—what I will call *phenomenal avowals*—comprise examples like “I have a headache”, “My feet are sore”, “I am tired”, “I feel elated”, “My vision is blurred”, “My ears are ringing”, “I feel sick” and so on.<sup>9</sup>

And of this group I wrote that:

... they are *strongly authoritative*. If somebody understands such a claim, and is supposed to make it sincerely about themselves, that's a guarantee of the truth of what they say. A doubt about such a claim has to be a doubt about the sincerity or the understanding of its author. Since we standardly credit any interlocutor, absent evidence to the contrary, with sincerity and understanding, it follows that a subject's actually making such a claim about themselves is a criterion for the correctness of the corresponding third-personal claim made by someone else: my avowal that I am in pain must be accepted by others, on pain of incompetence, as a ground for the belief that I am.<sup>10</sup>

I contrasted phenomenal avowals with *attitudinal avowals*, illustrated by examples like "I believe that term ends on the 27<sup>th</sup>", "I hope that noise stops soon", "I think that professional philosophers are some of the most fortunate people on earth", "I am frightened of that dog", "I am thinking of my mother", where the states of mind concerned are partially individuated by the propositional content, or intentional direction, which they contain. About this group I remarked, first, that their characteristic linguistic expressions may be, and quite often are, used not to make avowals, properly so regarded, but rather to affirm opinions about oneself arrived at by means broadly coincident with those that others also use to arrive about opinions about oneself—when, in other words, they express the products of exercises in self-interpretation; and second, that even when used as avowals, to express distinctively first-personal self-knowledge, they are at most, in contrast with phenomenal avowals, merely *weakly authoritative*, where weak authority involves

---

<sup>9</sup> P. 321

<sup>10</sup> Ibid.



that besides insincerity and failures of understanding, a small range of additional categories are open whereby falsehood may be explained, including paradigmatically self-deception, wishful thinking, and the like. My thought was that the presence of these additional categories nevertheless leaves space for a kind of authority, since there is again a standard presumption that subjects are not self-deceived, or thinking wishfully, etc.

Although this division is no doubt crude, it does at least serve to make the important point that the details of privileged access may vary with the type of psychological state avowed. But Snowdon is— I think justifiably,—dissatisfied with it on a number of counts. First, there are cases that do not seem to be comfortably allocated to either category. For example, what about first-person perceptual reports like, “It looks to me as if there is a ship on the horizon”? Here the avowed state is partially individuated by content, so by the terms of Wright [2001] its avowal should count as attitudinal; but the avowal also reports a characteristic kind of visual experience. So it seems to have a case for being placed both as phenomenal *and* as attitudinal. Then there are avowals of mood like, “I am depressed”, or “I am calm”, that convey no specific propositional content, or intentional direction, but nor are they comfortably viewed as reporting a distinctive kind of experience or feeling either. And what about avowals of action like, “I am heading home”? That too seems unhappily assimilated either to “My vision is blurred” or “I believe that term ends on the 27th”, for it reports neither any distinctive experience or feeling nor a state with a characteristic content or direction upon an object.

It would be pointless to respond by proposing that all avowals should count as phenomenal which are not individuated by intentional content, reserving the remainder as attitudinal. That will classify “It looks to me as if there is a ship on the horizon” as attitudinal, unhappily as before, since it both reports an experience and expresses no particular attitude. And some of the cases that will thereby be classed as attitudinal on the grounds that they involve intentional direction—for instance “I am thinking of my mother” and “I am exasperated by that noise”—may well also express certain distinctive kinds of experience, whereas other attitudinals, like “I expect that noise will stop soon” and “I believe that term ends on the 27<sup>th</sup>” need not do so. And in any case it may seem theoretically uncomfortable in the first place to adopt a categorization that slides over the difference between content-involvement and object-direction.

Nor is it plausible in general that the avowals which get classified as attitudinal in the terms of Wright [2001] are one and all candidates for expressions of the conclusions of processes of self-interpretation or, relatedly, liable in certain contexts to defeasibility on grounds of self-deception. Perhaps “I am thinking of my mother” might, in some bizarre contexts, fit that bill. But surely not “It seems to me that there is a ship on the horizon” or “I am watching that snake”. On

the other hand, and conversely, some reports, like “I am ashamed” or “I am elated”, which Wright [2001] would classify as phenomenal, might well be arrived at, as it were, third personally, as conclusions to explain one's behaviour—(“I realised, with his skilful prompting, that the discomfort I was feeling was actually that of shame.”)— and may also give expression to self-deception—(“You don't really feel any shame. You are merely uncomfortable at having been found out.”)

So, although I don't think that in Wright [2001] I was under any illusion that there were important differences within examples of the two overarching categories that I there proposed, it is clear that the phenomenal-attitudinal taxonomy needs re-thinking. Snowdon's view is that it is unlikely that any successful account of self-knowledge will be indifferent to the distinctions that he notes within the category I originally classified as "attitudinal". I am inclined to agree with that, although I would suggest that the contrast between what we might regard as limiting cases of the two original categories—pure expressions of attitude like, in a typical context, “I believe that water is wet”, unmarked by any characteristic phenomenological colour or feeling, on the one hand and pure expressions of personal experience, like “I have an intense headache”—remains crucial. But what should be made of the complications and cross-cuttings among kinds that Snowdon's discussion reminds us of?

Snowdon is cautious about suggesting any new overall taxonomy.<sup>11</sup> But, dispensing with the notion of avowals, he is happy to characterise as 'phenomenal' those *judgements* that

“concern the way experiences are, irrespective of what they indicate or seem to mean. Thus if you characterize how you feel as being ‘as if my leg is bent’ you characterize your experience in terms of what it indicates is the case (though without committing yourself to its being that way), whereas if you characterize what you feel as simply a pain or an itch that is not in terms of what the experience indicates or means. I take it therefore that descriptions of your after images as, for

---

<sup>11</sup> He does make the tentative suggestion that one epistemologically important distinction is that between self-knowledge of what he calls *experience states* and self knowledge that is not of such states—with the later encompassing but not exhausted by one's knowledge of one's propositional attitudes. And he suggests, “the reason that this twofold distinction seems significant is that self knowledge of non-experience-states is not in any plausible way to be treated as grounded in or derived from what might be called inner observation or inner-scrutiny.” (Snowdon [2012] p.249) That seems right—and indeed close to the spirit of my original distinction—though it broadens the phenomenal wing to include avowals of experience states whether these are content bearing or intentionally directed or not.

One may well wonder of course what for these purposes is covered by Snowdon's notion of an “experience state”. Some readers may feel that they understand well enough the kind of thing that he has in mind: an experience state is one with a distinctive phenomenology, a state such that “there is something that it is like” to be in it. I myself don't think that familiar paraphrastic slogan clarifies very much. (After all there is something it is like to be a stone — it is to be smallish, grey maybe, composed of igneous material, etc.) But I certainly don't want to maintain that one should be sceptical about the distinction for which it gropes.

example, blurry or red should count as phenomenal. (Wright’s list does not mention such cases.)”<sup>12</sup>

Snowdon selects this class of judgements as the primary focus of his subsequent critical examination of the alleged marks of privileged access, partly as he says for the not very interesting reason that much post-Wittgensteinian discussion of self knowledge has inherited Wittgenstein's tendency to scrutiny of examples of this kind in the *Philosophical Investigations* discussion of privacy, but much more interestingly because it

. . . may further seem that this scrutiny can itself be justified. What has often seemed to be the case to philosophers is that such self knowledge (of the phenomenal) is, in some way, what one might call especially good knowledge. We are particularly well placed to get such knowledge. Now, it is hard not to think that if, on reflection, there is nothing, as one might put it, special at all about so called phenomenal self-knowledge (and belief), then there is going to be nothing special about psychological self-knowledge in general. If that conditional claim is plausible, [it is tempting] to see this sort of phenomenal self-knowledge is a crucial test case.<sup>13</sup>

Snowdon expresses, I think wisely, some caution about the conditional claim. But it remains that his goal appears to be to argue that there is, perhaps, “nothing special”, or nothing obviously special, about phenomenal self knowledge (and belief) so characterised. I will turn shortly to the question how well his reflections tend to support this deflationary tendency.

My own reaction to the complications and cross-cuttings is that an adequate taxonomy of the permissible subject-matter of avowals will have to be much more complex than anything yet envisaged. We do seemingly have distinctive first personal knowledge of what our current experiences are like—our “experience states”<sup>14</sup>—and of our propositional attitudes and their contents. But, as remarked at the start, we also seemingly have such knowledge of what we are currently doing, of how we feel (within parameters of mood and emotion), of what we seem to be perceiving, and of what our thoughts and emotions are intentionally directed upon. These aspects may be multiply present in a given state of mind. That is, any given state of mind may involve some or all of

A distinctive phenomenology or “raw feel”

A distinctive content (that P)

A distinctive attitude (belief, desire, fear, hope...)

---

<sup>12</sup> Snowdon [2012], p. 248.

<sup>13</sup> Ibid.

<sup>14</sup> See footnote 11.

A distinctive direction (of the state's being *of*, or *at* so-and-so)

A distinctive emotional affect

A sense of agency

A sense of perceptual, or memorial reception

A sense of mental causation

- and perhaps much more. One very striking aspect of the kind of self knowledge in which we are interested—the especially effortless, immediate, reliable and comprehensive kind (however exactly these features should be characterized)—is that it not merely often embraces these aspects of our mental states but also (though care is needed with this) that it is normally seemingly closed over their adjunction in a single state. Thus I can properly avow, for example, something as complex as that “The intensity of my toothache and the realization how long I had neglected it made me especially apprehensive at the approach of the long unexperienced but nevertheless all too familiar high pitched whine of the drill...” The prototypical raw feels and the phenomenologically colourless attitudes-cum-contents that motivated my original phenomenal/attitudinal contrast are just two simple, radically contrasted possibilities of combination among this larger range of parameters. The “data” which a philosophical account of self knowledge is charged to explain are indeed much more complex than is suggested by my earlier discussion. But they are also more complex than is suggested by Snowdon’s discussion. Were it indeed to prove that there is “nothing special” about phenomenal self-knowledge, it should be clear that it would still be an egregious *non sequitur* to conclude that there was nothing special about self-knowledge anywhere. A spreading variety of cases need to be reviewed individually and in detail.

So, a philosopher who wants to solve, or dissolve, "the problem of self-knowledge" confronts the following series of tasks. First, one of two cases must be argued for: *either* a case that it is simply an illusion that any of the aspects of our mental states listed above, or any other aspects of them, exhibit anything close to the trio of marks of privileged access distinguished; *or* a case that that all (or some) of the listed aspects of our mental states do indeed (characteristically, or often) fall within the purview of privileged access, perhaps characterised as I have suggested, or perhaps in somewhat other terms —although I do not think we will recognise the problem if presented in terms of a notion of privileged access that is too different from what I have proposed. In the latter case, the next task will be carefully to characterise in all instances where there is privileged access, what if any local variations in its character are manifest (witness, for example, my own earlier somewhat obtuse contrast between strong and weak authority.) And finally, when the 'data' are thus placed in clear view, the task will be *either* to debunk the project

of a philosophical explanation of them *or* to offer philosophical insight into *why* privileged access, whether epistemically or linguistically characterised, holds sway over our own mental states but stops short of others': what it is about the subject matter of avowals, or our relationship to that subject matter, or their broad function in our rational economy, or their semantic or pragmatic function, or . . . (and here I show that, to me at least, it is uncertain what are the admissible categories of distinctively *philosophical* explanation that one might invoke to serve this project)— what it is that, once brought to our attention, serves to make sense of privileged access without unacceptable additional philosophical costs.

I do not think that anything that philosophy has so far achieved comes anywhere near to a satisfactory accomplishment of this project.

§6. Let us restrict our attention from this point forward to avowals of pure experiential phenomenology, as it were—those expressive of Snowdon's category of phenomenal judgement, typified by the self-ascription of pains, itches, after-images, and so on, and encompassing just characteristics of felt experience dissociated from any form of content or intentional direction<sup>15</sup> — and turn to Snowdon's detailed challenges to the broad, three-element account of privileged access that informed my earlier discussions. The three elements, once again, specialised to this class of avowals, were these:

*Immediacy*: I wrote that “The demand that somebody produces reasons or corroborating evidence for such a claim about themselves—“How can you tell?”— is always inappropriate.”<sup>16</sup>

*Strong authority*—“A doubt about such a claim has to be a doubt about the sincerity or the understanding of its author.”<sup>17</sup>

*Saliency*—“Where P is an avowal of the type concerned, there is typically something absurd about a profession of the form, ‘I don't know whether P’ ”.<sup>18</sup>

Snowdon charges that this class of avowals—I will continue with the terminology of "phenomenal avowals"—are not in fact characterized by these features. Before reviewing his reasons for that claim, let me quickly note a point about its significance.

---

<sup>15</sup> Of course, the claim that there are any such avowals is controversial, and contested for example by Michael Tye. See e.g his [1995].

<sup>16</sup> Wright [2001], p. 321.

<sup>17</sup> Ibid.

<sup>18</sup> Ibid.

If one's aim is to get the data about self knowledge right, and then either to provide an explanation of them (or to argue convincingly that none is needed, or anyway is going to be forthcoming), then of course it's vital that the features proposed be indeed characteristic of the target phenomena. But there is a Wittgensteinian project nearby for which strict accuracy is not required. That is the project of *diagnosis* of the grip exerted on us by bad pictures of the mind and of our relationship to our mental states. For this diagnostic project, it is necessary in order to explain the attraction of such a bad picture only that avowals prove to have certain *apparent* similarities to statements of some other kind where a certain kind of model of the subject matter, and/or of our cognitive relationship to it, is appropriate enough. That will be enough to explain the attraction of a corresponding picture of the mental. Seduction by such a model does not require there be an exact or even a close fit between the relevant aspects of the two discourses. It suffices that there be enough of an appearance of fit to lure in unsuspecting ordinary thought and cause it to erect a bad picture based on a spurious analogy. Of course there is a further step, in Wittgenstein's own approach, to the conclusion that no "theory", that is, picture or model of our cognitive relationship to the relevant subject matter, sufficing to explain the distinctive aspects of the discourse, should be sought, or indeed is to be had. Wittgenstein's quietism, however, is not the topic of this paper.

In Snowdon's view, the three generalizations I suggested about our linguistic practice with phenomenal avowals are hasty and overgeneralised. Perhaps they are acceptable for severe pain and a few other central cases but, he argues, they are hardly assertible straight off without a detailed review of cases. And in fact, he suggests, each seems to be subject to counterexamples when we take a wider look round.

For the case of strong authority, consider the victim of a road accident who has previously taken a substantial quantity of amphetamines. He screams and claims he is in severe pain as the doctor examines him, although the doctor can find no sign of significant physical injury. We might want to say, surely, that his screaming and impression that he is in pain is a product of fear and confusion and intoxication, and in that case we could be inclined to discount what he says without regarding him either as insincere or as guilty of some linguistic misunderstanding. More qualifiedly we might anyway be inclined to discount his avowals without thereby feeling committed to interpreting him as either insincere or as misunderstanding what he is saying, although venturing no clear alternative explanation, and feeling no need to do so.

As against salience, suppose you are having your eyes tested for the suitability of a new pair of glasses. The optometrist asks you, "Which of these lenses gives a better focus?" You may

very possibly not find it easy to say, when basing your judgement just on the phenomenal appearances. Nevertheless there may be a fact of the matter that emerges when, for instance, you attempt to read print in different font sizes using the two lenses. Or consider “My ears are ringing”. Is it absurd to express ignorance in the form of the remark, “I do not know whether my ears are ringing or not”? Not at all—perhaps what you are hearing is an effect of the circulation of coolant in the central heating system, or of distant machinery.

Finally on groundlessness/immediacy, Snowdon opines that my discussion confused the inappropriateness of asking for evidence with the inappropriateness of the question, “How do you know”? Indeed he suggests that the appropriateness of the latter must go with the very idea of knowledge in the first place. (I will come back to this later.)

§7. These observations give pause. However, as noted, Snowdon is in any case unhappy about conducting the examination at the level of discourse about the mind rather than at the level of the mental subject matter itself. His counter-suggestion is that insofar as there is any plausibility in ascribing the three listed features to avowals, it is grounded in that of certain conditionals which concern *the beliefs* that avowals express. Thus: if it is inappropriate to ask for the evidence upon which an avowal is based (groundlessness), that can only be because the belief (of knowledge) to which an avowal gives expression it is not itself based on evidence. And if a doubt about an avowal must commit one to doubt about (either) the author's sincerity or his understanding of the language, (strong authority), that, Snowdon suggests, can only be because when an avowal does give expression to a belief, that belief must be right. (Snowdon terms this *incorrigibility*.) Finally if, where P would express a pure phenomenal avowal, one cannot ever appropriately say “I do not know whether P or not”, that can only be because its not possible for P to be the case and the subject not realize it. (Snowdon calls this *self-intimatingness*.)

Snowdon charges, however, that both incorrigibility and self-intimatingness fail even for the beliefs expressed by pure phenomenal avowals. So restricted, self-intimatingness may, as a first approximation, be expressed by the conditional:

Necessarily, if any purely phenomenal condition C applies to S, then S knows that he/she is in C.

Snowdon quickly observes that this formulation is routinely counter-exemplified by animals, infants and demented adults, all of whom, we conceive, may be in determinate phenomenal states without the attendance of any appropriate knowledge. He moves to the obvious modification:

Necessarily, if phenomenal condition C applies to S, and S possesses the relevant concepts, then S will know that C applies to him/her.

But Snowdon finds two general problems even with this. First, he suggests, there are presumably some phenomenal properties whose detection requires implementation of some kind of procedure. Suppose the question is, for example, whether a particular after-image that I am experiencing has twelve spots on it or not. To find that out, I will need to attend to it over a period of time and count up. So surely I can sensibly say when the question is raised but before I count up, “Well, I do not know whether my after-image has twelve spots on it or not”. That is an apparent counterexample to self-intimatingness as re-formulated.

There is, though, a natural worry about this kind of apparent counterexample. It is that when the answer to a question about an after-image, or any phenomenal state, presents as verifiable only by a process of temporally extended attention, we have to presuppose that we can place independent trust in the *stability* of the image, or state, through the requisite process. What criterion do we have for such stability? If I do attend to the image and count up the spots, what reason can I give for thinking that their number has been invariant through the process of counting? By contrast, there are all kinds of ways one might get reason to think that an external object — say a piece of paper with some number of dots scattered on it — is relevantly stable. Suppose I count up the spots on the after-image twice, getting different results. On what might I justifiably base the conclusion that the image has changed, rather than that I have miscounted, or vice versa?

The point here is not a sceptical one — as if one or the other account might be correct beyond my ken, as it were. Rather, it is that the whole idea of attending to and investigating a phenomenal object over a period of time in effect treats it on the model of an external object presented to attention, akin to a kind of percept. And such a model demands that distinctions make sense which we are not empowered to draw. So while the model does indeed, as Snowdon in effect is observing, put the self-intimatingness of the phenomenal in jeopardy, — at least as self-intimatingness is formulated above, — it begs a large philosophical question about the nature of introspective awareness and its objects.

Snowdon’s second objection to self-intimatingness is less tendentious. Suppose we make

. . . the general naturalistic assumptions that the occurrence of the experience requires a physical realization, and that the occurrence of the extra uptake or recognition does so as well. Now, this second physical occurrence must presumably be causally related to the physical occurrences which constitute the experience itself. Finally, understood this way, the physical causal processes [sic] involved in the uptake is such that it need not occur of necessity. In principle, something could



happen to stop it occurring. It seems to follow that the knowledge of the experience need not occur. Experiences are not self-intimating.<sup>19</sup>

This thought can seem compelling. But a friend of self-intimatingness might doubt it actually takes us beyond the point already made by the possibility of animal, infant and demented adult experience. Granted, if self-knowledge, even in purely phenomenal cases, is a genuine cognitive achievement, then it must consist in something over and above the occurrence of the state known and there will presumably be a causal connection between the occurrence of that state and the attendant knowledge. But we already moved to a complication of the original formulation of self-intimatingness to include the additional condition in the antecedent that the subject possesses the relevant concepts. So where is the sting? Well, what is true, presumably, is that possession of concepts is one thing and *their successful exercise on a particular occasion* is something further—there may be other conditions that frustrate their successful exercise which may obtain alongside the occurrence of the appropriate experiential state. Amphetamine-induced panic, for example, reverting to Snowdon’s example, might be such a condition. Surely, indeed, there is a whole range of common sense possibilities of ignorance here—a simple lack of attention owing to “panic, fear, drug induced jitteriness”, or “an inability to take the thing in accurately owing to being primed to expect something else...”<sup>20</sup> Indeed, Snowdon suggests, even a case of severe pain need not be self-intimating. If S has a severe pain in his left foot and has the appropriate concepts, must he be aware that he has? What about cases of overload, where S has multiple injuries and complains in the first instance of pain elsewhere. The attendant medic asks him, with one eye on the bruising and swelling on his ankle, “What about your foot”. He replies “That’s OK—no, wait, no it isn’t OK; in fact its extremely painful.” Isn’t it natural to say in such a case that the pains felt elsewhere in his body have *distracted* the subject from his painful foot, at least temporarily?

Well, as before, there is a metaphysical issue here about the stability and persistence of the state. What objection would there be to saying that the effect of the other pains was as an at least partial *anaesthetic* and that the anaesthesia becomes less effective when the subject directs his attention to his foot? Again, there is a worrying indeterminacy.<sup>21</sup> However I think it clear that

---

<sup>19</sup> Snowdon [2012], p. 257.

<sup>20</sup> Snowdon [2012], p. 258.

<sup>21</sup> Snowdon (ibid) brings another more theoretical doubt to bear: the suggestion that occurrences of a phenomenal state might, in a suitably (temporarily) impaired subject, fail to trigger any application of “the self concept” — so that he might not be moved by such an occurrence to make a judgement that he himself is in pain, even though possessed with the concept of himself and that of pain. (A classic fictional case is that of Mrs. Gradgrind in Charles Dickens’ *Hard Times*.)

Snowdon's discussion does bring out that transparency, or in his terms self-intimatingness, formulated as above, is still at best a crude formulation, and stands in need of qualification. The single most important qualification that is suggested—one that perhaps, indeed, corrals all Snowdon's reservations about self-intimatingness—is that, irrespective of subject matter, and whatever difficulties there may be with the idea of a state of that subject matter persisting unchanged, there are certain intrinsic difficulties in the idea that the mere occurrence of a situation can provide a metaphysical guarantee of an appropriate judgement about it in any subject, however advantageously he/or she may be placed. The point is obvious enough if we reflect on the situation of perceptual judgement made under unimprovable external conditions by a thinker whose sense organs are functioning impeccably. There must still be scope for the possibility that no such appropriate judgement is made, simply because some additional condition obstructs the exercise of the appropriate conceptual and judgemental capacities. Snowdon's catalogue of possibilities—distraction, panic, fear, drug induced jitteriness, etc.—are all cases where such an interference with judgemental capacity operates. Surely this point must hold both for the perceptible *and* the phenomenal, so long as the latter is conceived as a potential object of genuine knowledgeable judgement at all.

How far away, though, from the idea that "there is something special" about phenomenal self-knowledge does the concession take us? Maybe it's impossible to circumscribe all such possibilities in a finite formulation and write it into a complication of the self-intimatingness conditional above in such a way as to ensure a conceptually necessary truth. But something like the following may yet be a schematic approximation to such a truth:

Necessarily, if phenomenal condition C applies to S and S possesses the relevant concepts and is in no condition that would impair their exercise in judgement, then S will know that C applies to him/her.

This formulation is, to be sure, bland and non-specific. But it's not trivial. It is failed, for example, by all judgements of the obtaining of non-phenomenal conditions based on perception and memory. It is also failed by judgements of all kinds involving inferential processing and judgemental skills going well beyond those involved in normal possession of the relevant concepts. It is of interest to consider for what range of psychological conditions an analogue might hold good. But, at least for phenomenal states whose appraisal does not require extended attention, it impresses as true.

§8. Snowdon formulates *incorrigibility* like this:

Necessarily, if S believes that he/she is in C, then S is in C

Some of the objections that he runs against this principle are in effect reapplications of points made in connection with self-intimatingness. But there are some differences too. Snowdon observes that if enough is at stake, we may want to revisit certain self-ascriptions of sensation. A good example is given by CB Martin's medical diagnosis case:

Imagine you are having troubling pains in your stomach. The doctor after an examination says, 'It is one of two complaints. There is a test to determine which. I shall apply pressure to a region and it will cause pain in two areas—one left and one right—and whichever hurts most tells us what the problem is.' He then applies the pressure and you judge that the left area hurts most. The doctor then says, 'Oh dear, that means that you have a very serious condition, requiring drastic treatment. If it had been the right side then it would have been minor.' Confronted with this, do you say, 'Since my judgement is incorrigible go ahead with the drastic treatment.' Or might you say, 'do you mind if we repeat the test? Maybe I was a little bit hasty—I would like to make sure and check things.' The final claim is that we would indeed opt for the second response, and not the first. We would not view anyone doing so as overlooking some basic principle that we cannot make mistakes in these sorts of judgements.<sup>22</sup>

Actually the last claim is doubtful in full generality. Suppose we change the example slightly and make the question at issue not which of two concurrent sensations is the more painful but whether e.g. a single sensation is experienced on the left of the abdomen rather than the right. You might still want to run the test again, but you would probably do so with a sense of irrationality and a sinking feeling, in tension with the idea that you would merely be checking whether the original judgement might simply have been mistaken. So the crucial feature implicitly claimed for the original example — that it would not be irrational to run the test again — is trading on a tacit understanding that the relevant description of the sensation may be one whose application might *anyway* occasion some difficulty, even in less emotionally fraught circumstances. But neither incorrigibility nor self-intimatingness precludes the possibility that the phenomenal should allow for vagueness and 'hard cases'. And there is in any case a question whether, if the patient does run the test a second time, and wishes to return a different verdict, that is necessarily evidence that the first assessment was mistaken. Would not such a test anyway be likely to be formulated in terms of a dependable *regularity* of judgement?

---

<sup>22</sup> Snowdon [2012], p. 259.

Snowdon repeats, under the heading of incorrigibility, his points about cases where some extended investigative procedure is required, and about the range of common sense possibilities—panic, influence by drugs, etc.—that disrupt judgement. “Suppose” he writes, “that someone is extremely frightened and nervous at the prospect of torture by having acid dropped on their skin, and to the accompaniment of misleading noises, etc., is very suddenly doused in water. It is highly likely they will be under the impression initially that the acid torture had started, and along with it the horrible pain”. Well, yes indeed. But the point this makes is only, as before, that a robust formulation of incorrigibility, if such a thing is possible, would have to exclude conditions where the subjects capacity for judgement is, for whatever reason, seriously impaired.

Snowdon also brings up the experiments at Harvard where people thought they were experiencing visual images when in fact they were seeing patterns projected on to a wall. The example stands comparison with the ringing-in-the-ears case mentioned earlier. But I think it is fair to say that these cases should be dismissed as irrelevant. A genuinely phenomenal judgement should not be interpreted so as to exclude a real external percept; it should be neutral on that. The interpretation of “My ears are ringing” so as to be defeasible by the discovery that actually one is hearing an external sound violates that principle.

So the moral, I suggest, is much as it was for the discussion of self-intimatingness. Snowdon’s formulations of both principles are indeed put in difficulty by the considerations that he advances. But the right conclusion is not that incorrigibility and self-intimatingness of phenomenal judgement are quite misguided contentions, but that they have been too crudely formulated. The conditions for success for phenomenal judgement, like the conditions for judgement in general, have to incorporate provisos to ensure that the subject is not in some way temporarily disabled from exercising his judgemental competences. All of the states of the subject, and types of external circumstance that Snowdon marshals to make trouble for his formulations of the two principles are such that their obtaining would pose obstacles to a subject’s exercise of *any* judgemental competence, about any subject matter.

It may be replied that that is the whole point: that Snowdon is after all trying to support the contention that there is nothing special about self knowledge—no special security, either in the direction of self-intimatingness or in the direction of incorrigibility applies. But to establish that of course requires more: it requires argument that there are no *other* conditions which, while their obtaining may impair judgements about other subject matters, exert no deleterious impact on judgements concerning the phenomenal states of oneself. And that isn’t so, not even if we restrict attention to *non-inferential* judgements of other kinds. No counterpart of the refined thesis of self-intimatingness, nor of a corresponding refinement of incorrigibility, is remotely plausible

when the subject matter of the judgement concerned is taken to be not the subject's phenomenal states but e.g. states in the past within his memory, or the states of his local perceptible environment.

The explanation of that has to do with the proper understanding of the notion of the immediacy of phenomenal judgement, to which I now turn.

§9. There is quite a bit of tidying up to do here, for Snowdon is, I think, quite right that what I have said in earlier work about groundlessness (immediacy) encourages (and betrays) confusion. There are a number of relevant but distinct notions. In Wright [2001], I primarily intended a point of analogy with observational judgements, so that “groundless” did not mean made without support, or “out of the blue”, but rather: lacking articulatable, independently appreciable reasons. In this sense, observational judgements are groundless too. It is true that, if charged to supply grounds for my thinking that my screensaver just clicked on, I may reply, “I saw it do so”. But that is not to offer grounds for my judgement so much as to explain why, as I think, I do not need to. To claim perceptual awareness that P is not to give one's reasons for believing P. That I take myself to be perceptually aware that P is not the reason why I believe it. Rather that my experience is shaped as of P's being so leads me both to believe that P and to believe that I am perceptually aware that P.

Still, though both may be groundless in this sense, there is a crucial difference between the phenomenal and perceptual cases. When P is a perceptual judgement, I could normally have had, and may actually have, *other* types of ground—indirect evidence, testimony, memory, and so on— for the same judgement. And in a case where I am perceptually unaware that P, I might *only* have such indirect grounds. But if P is a phenomenal judgement, there seems to be no sense in the idea of being in a position where I could rely only on such auxiliary grounds—if I were in such a position, they would be auxiliary grounds for something false. There is nothing for them to do if the very basis in awareness for the truth of P itself, viz, the relevant phenomenal state, is missing.

Thus, while it would be a perfectly respectable terminological proposal to identify the ‘immediacy’ (or ‘groundlessness’) of a judgement with its non-inferentiality, or (not exactly the same thing) with its not being based on reasons—so that observational judgements, and phenomenal judgements would be alike immediate (at least if we take it that the basis for observational judgement is not inferential)— there is a deeper difference, and a perhaps more worthwhile sense of ‘immediate’, to distinguish the two kinds of judgement. The reason why the idea of reliance purely upon indirect grounds—of the sort on which others have to rely—for one's

being in a certain phenomenal state is absurd is because there is, crudely, no faculty of apprehension that mediates one's awareness of one's phenomenal states as sense perception mediates one's awareness of the disposition of objects around one, or memory mediates one's awareness of one's own past. There is no mode of awareness that gives one *access* to one's phenomenal states as sense-perception, and memory, give one to access to the disposition of material objects and one's past. Rather phenomenal states *are* states of awareness.

This contrasts in turn with a third possible sense of 'immediate', in accordance with which a mental state would count as immediately available to a subject only when lacking in any distinctive phenomenology, any distinctive "what it is like" to be in it. In this sense, one's awareness of an occurrent belief might be unmediated by any distinctive state of consciousness and in this respect would contrast with one's awareness of a headache, or a ringing in one's ears. But I will not further pursue this third notion here.

How do any of these distinctions impact on the propriety of my earlier gloss on immediacy, rejected by Snowdon, that "the question, how do you know, makes no sense"? In Snowdon's view, if there really is such a thing as knowledgeable phenomenal judgement, then there has to be such a thing as *the way one knows*. Naturalistically viewed, knowledge always comes to one somehow, so there must be an answer to the "how" question. And he is surely right that there is no commitment to thinking that just because there is an answer to this question, a subject who possesses the knowledge must be in a position to provide the answer. Still, if knowledge does indeed require that there be a correct answer to the question: how do you know?, then there is an issue whether someone who has no inkling of what that answer might be is after all in position to *claim* knowledge in the appropriate circumstances. And surely we are, when in pain, in position to claim to know that we are. So we have a puzzle.

What is available by way of a plausible answer to the question, how someone might know that they are in pain? Snowdon himself rejects two bad-seeming answers. The first is that "You know you are in pain by feeling the pain". That seems bad because being in pain *is* feeling the pain and "You know you are in pain by being in pain" doesn't seem terribly good; after all, we wanted to clarify *the connection* —we wanted to explain how the subject, as it were, keeps track of the pain. Snowdon's second rejected answer is "You know you are in pain by introspection". For what is *introspection*? The etymology suggests some conception of an "inner eye". But, as Snowdon observes, if a creature has pains, it doesn't need any additional quasi-perceptual mechanism, if that is what introspection is, to be aware of them.

I think the right answer is the first: that if Snowdon's principle, that where there is knowledge, there must be a "how", is accepted, then the proper answer to the question, "How do

you know that you have a headache?”, is indeed: by having the headache (and exercising routine judgemental competences.) This seems bad only because the corresponding answer in the case of perceptual knowledge— an answer that just cites the (perceptible) fact that P —is obviously bad. In the perceptual case, an account of the “how” needs to say something about the sensory mechanisms that enable one to detect the fact that P. But if the states that provide the objects of phenomenal judgement are already states of awareness, then there is no need for detection. And all that remains to do is to articulate one's states of awareness in judgement.

Immediacy, then, in the second sense canvassed above, whereby non-inferential observational judgements still count as mediated rather than immediate, is not at odds with the significance of the question, “How do you know”. I was wrong to suggest that there is any useful sense of groundlessness, or immediacy, associated with the senseless of that question. Rather, what immediacy, so understood, is at odds with is any kind of answer to it that makes a comparison with ordinary inferential or non-inferential judgement. When I am in pain, I am able to know that I am precisely because I am in pain and pain states are immediately accessible states.

§10. Let me venture some tentative conclusions.

First, Snowdon’s examples serve as an effective reminder that all judgement draws on certain capacities — of focus, attention, and general cognitive lucidity — that are not assured simply by the possession of relevant concepts, and hence that any satisfactory elucidation of the dependencies that the ideas of authority and salience gesture at will have to incorporate provisos that recognise this. This point does not, however, have any tendency to establish that “there is nothing special about self-knowledge”. What is special is that a whole range of *other* factors that, in the case of other kinds of knowledge, and in particular other kinds of non-inferential knowledge, can lead to mistakes and ignorance, are inoperative where phenomenal self-knowledge is concerned— and indeed, one might suggest, where any kind of self-knowledge is concerned which may properly generate an avowal. Suitably qualified, and *pace* Snowdon, immediacy, authority and salience remain as useful headings in terms of which to try to plot the distinctiveness of psychological self-knowledge, and so properly to articulate the problem which it poses.

Second, one prime source of such “other factors” is located in *cognitive interaction*— in the causality involved in perception and memory, for instance, between the state of affairs known about and the knowing subject. Where knowledge depends on such interaction, the needed interactive process may be variously impaired, or even non-existent. But the immediacy of phenomenal states, their essence as states of awareness, closes out all space for such interaction

and, with that, pre-empts all scope for the operation of the kind of other factor that, in cases where cognitive interaction is involved, may be responsible for mismatches between opinion and fact.

Third, that reflection indicates that in the case of phenomenal judgement, authority and salience are *sourced* in immediacy. It is because a subject's awareness of her own phenomenal states is unmediated by any interactive cognitive capacity that her impressions of them stop short of absolute authority only modulo conceptual shortcomings and the other kinds of judgemental failings whose possibility Snowdon emphasises. As it is for the same reason that although not absolutely salient— not such as to guarantee knowledge whenever they occur — the possibilities for a relevantly conceptually competent subject's phenomenal states to elude her recognition are narrow and strained, and involve factors that tend to call into (temporary) question her fitness for judgement generally.

Others of our mental states and attributes that fall within the scope of avowal, however, are not immediate in the relevant sense: they need be nothing in my consciousness, in the sense in which toothaches, and tickles, and tinnitus, are elements of consciousness that constitutes the occurrence of an intentional state, or mood, e.g. So it follows, fourth, that insofar as authority and salience, or close relatives of them, apply in cases of non-phenomenal self-knowledge, their explanation will not prioritise immediacy as it applies to the phenomenal but will have to proceed along other lines. This observation belongs with what I suspect may be the more general insight that there are very likely to be important local variations amongst psychological states and attributes in the exact forms in which immediacy, authority and salience are instantiated, and corresponding variations in their explanation (should they allow of explanation at all.)

Finally, for anyone inclined to take the 'linguistic turn' — inclined to suppose that the most illuminating approach to the problem of self-knowledge will indeed involve concentration first and foremost on the language game of self- and other-ascription of psychological properties — we have found no serious cause in Snowdon's discussion for loss of confidence. And the same goes for sympathisers with Wittgenstein's anti-explanatory deflationism. The feasibility of the explanatory project, and the correctness of Wittgenstein's defeatist assessment of it, both remain live possibilities.<sup>23</sup>

---

<sup>23</sup> Grateful thanks for comments and criticisms to participants at the *Self and Self-Knowledge* conference held at the Institute of philosophy in London in 2008, at which a prototype of this material was presented, to members of the pilot project on self-knowledge that ran at the Northern Institute of Philosophy from 2009-12, to Sandy Goldberg, and especially to Paul Snowdon.



---

*References*

- Bar-On, Dorit [2004]: *Speaking My Mind: Expression and Self-Knowledge*, Oxford: Oxford University Press.
- Blackburn, Simon [1993]: “Gibbard on Normative Logic”, *Philosophical Issues* 4, *Naturalism and Normativity*, pp. 60-66.
- Burge, Tyler [1979]: “Individualism and the Mental,” *Midwest Studies in Philosophy* 4, pp. 73–121; reprinted in Burge, *Foundations of Mind: Philosophical Essays, Volume 2*, Oxford: Oxford University Press, 2007, pp. 100-150.
- Finkelstein, David (2003): *Expression and the Inner*, Cambridge, Mass.: Harvard University Press.
- McDowell, John [1991]: “Intentionality and Interiority in Wittgenstein” in K.Puhl, ed., *Meaning Scepticism*, Berlin and New York: de Gruyter, pp. 148-169; reprinted as chapter 14 of McDowell, *Mind, Value and Reality*, Cambridge, Mass.: Harvard University Press 1998, pp. 297-321.
- McKinsey, Michael [1991]: “Anti-Individualism and Privileged Access”, *Analysis* 51, pp. 9-16.
- Putnam, Hilary [1975]: “The meaning of “meaning””, *Minnesota Studies in the Philosophy of Science* 7, pp. 131-193; reprinted in Putnam, *Mind, Language, and Reality: Philosophical Papers* (Volume 2), Cambridge: Cambridge University Press, pp. 215–271.
- Ryle, Gilbert [1963]: *The Concept of Mind*, London: Penguin.
- Snowdon, Paul [2012]: “How to Think about Phenomenal Self-Knowledge” in A. Coliva, ed., *The Self and Self-knowledge*, Oxford: Oxford University Press, pp. 243-262.
- Tye, Michael [1995]: “A Representational Theory of Pains and their Phenomenal Character” in *Philosophical Perspectives* 9, pp. 223-39.

Wittgenstein, Ludwig [1953] *Philosophical Investigations*, eds. G.E.M. Anscombe and R. Rhees, G.E.M. Anscombe (trans.), Oxford: Blackwell.

Wright, Crispin [2001] *Rails to Infinity: Essays on Themes from Wittgenstein's Philosophical Investigations*, Cambridge, Mass.: Harvard University Press.

—[2011]: “McKinsey One More Time,” in A. Hatzimoysis, ed., *Self-Knowledge*, Oxford: Oxford University Press, pp. 80–104.