

Replies to My Commentators

David H. Finkelstein

I am very grateful to Josep Corbí and Luis Valdés for inviting my participation in this symposium as well as at the 6th Inter-University Workshop on Mind, Art, and Morality. Both of them have been wonderful throughout this process. I am also grateful to the symposium contributors — Josep Corbí, Josep Prades, John McDowell, Barry Stroud, and Jesús Vega-Encabo — for their insightful and (mostly) sympathetic comments on *Expression and the Inner* (henceforth “E&I”). I feel tremendously lucky to have such an exceptional group of scholars engaging with my work, and I’m painfully aware of the fact that my replies leave many questions and criticisms unanswered.

I. REPLY TO JOHN MCDOWELL

As I read John McDowell’s remarks on E&I, he takes issue with me about two points, primarily. First, he thinks that I misread him as committed to what I call the “middle path account of inner awareness” (or “MPA”). Second, he suggests that both in my reading of Wittgenstein on sensations and in my own treatment of the issues with which E&I is concerned, I distort the shape of the real problems that bothered Wittgenstein and that ought to bother us — problems that, when fully understood, should strike us more as metaphysical than as epistemological. In what follows, I’ll say a little bit about each of these points in turn.

As McDowell notes, the “middle path account” is so called in E&I because it sits between detectivism and constitutivism, inheriting features from both (while neatly avoiding the problems that attend each). In his essay, McDowell complains about being read as defending this position. He says, “[T]here is no detectivist strand in the conception of the inner I meant to describe — no whiff of the idea that self-ascriptions report states of affairs that obtain anyway, independently of their subject’s ability to report them” [this issue, p. 16]. But this complaint is misdirected; in E&I, I was explicit about MPA’s not being detectivist in *this* sense. According to MPA, “[t]he pain in

my head is really nothing over and above my conceptual awareness of it" [E&I, p. 61]. Thus an avowal of pain is not, according to MPA, the report of some state of affairs that obtains independently of a subject's ability to self-ascribe it. How did I think that McDowell's position was akin to detectivism? According to MPA, a judgment that I myself am (e.g.) in pain requires epistemic justification if I'm to be entitled to it. And, according to MPA, this justification is much like the sort that attends perceptual reports. In reading McDowell as committed to this view, I was focusing on passages like the following from *Mind and World*:

To give the impressions of "inner sense" the right role in justifying judgements, we need to conceive them, like the impressions of "outer sense," as themselves already possessing conceptual content; to supply the necessary limit to the freedom of spontaneity, we need to insist that they are indeed impressions, products of receptivity [McDowell (1996), pp. 21-22].

According to the picture of self-awareness that McDowell presents here, just as conceptual impressions of "outer sense" are needed in order to justify the claims that we make on the basis of visual or auditory experience, so conceptual impressions of "inner sense" are required to justify the claims we make about, e.g., our own sensations.

In E&I, I understood these impressions of "inner sense" to be, according to the view set out in *Mind and World*, not only conceptual, but *propositional*. In his essay, McDowell seems to reject this reading. Immediately after noting that, in E&I, I read him as holding that "the experience of undergoing a sensation has propositional content" [this issue, p. 18], he continues:

But even in the case of perceptual experience of the environment, I have come to doubt that we should follow Sellars, as I used to, in conceiving experiences as having propositional content, the kind of content claims have. And certainly the idea of actualizations of conceptual capacities does not imply that just anything describable in those terms must have propositional content [ibid.].

McDowell notes here that he's come to doubt that we should conceive of the conceptual content of perceptual experiences as propositional. But his coming to doubt this marked a shift in his position, one that came after E&I was published. At the time of E&I, McDowell's view was (or seemed to be) that in order for an experience to justify a judgment, it must have propositional content. And, as the passage from *Mind and World* that I quoted in the preceding paragraph indicates, he held that impressions of "inner sense" do serve to justify judgments (about, e.g., whether the subject himself is in pain). So it seems to me that it was natural, anyway, to read him as I did.

McDowell (2009) writes:

I used to assume that to conceive experiences as actualizations of conceptual capacities, we would need to credit experiences with *propositional* content, the sort of content judgments have. And I used to assume that the content of an experience would need to include *everything* the experience enables its subject to know non-inferentially. But both of these assumptions now strike me as wrong [McDowell (2009), p. 258].

When I wrote E&I, I understood inner experiences, on McDowell's conception of them, to have both of these later-to-be-rejected features, and I argued that both were problematic. (Actually, my focus was much more on the second feature that he mentions in this passage than on the idea that the experience of a sensation has propositional content.) I should add that I don't see that the argument set out in Chapter 3 of E&I tells against McDowell's *current* view of inner experience. Indeed, I believe that at present, he and I are close together on these issues. But I should turn to the remainder of his essay, for there are still differences between us.

* * *

I said at the start of this reply that McDowell takes issue with me primarily on two points. Let's consider the second of these points, which I'll try to bring into view through the lens of PI §304. In §304, Wittgenstein's interlocutor accuses him of "again and again reach[ing] the conclusion that the sensation itself is a *nothing*." Wittgenstein eventually replies: "We have only rejected the grammar which tries to force itself on us here." McDowell suggests that, in E&I, I go wrong in thinking "the source of 'the grammar which tries to force itself on us here is the 'privileged-observation' picture of self-knowledge" [this issue, p. 20]. He continues, "This belongs with Finkelstein's Wright-like casting of detectivism as the primary villain in the philosophical drama" [ibid., p. 20]. McDowell believes that the primary villain in the drama that's enacted in and around PI §304 is a fantasy that we fall into if "we suppose that the presence of something in an inner world would have to be a reality independent of the linguistic and conceptual competence of the person whose world it is" [ibid., p. 21]. Once we're in the grip of this fantasy, "being reminded that expression is central to the 'grammar' of talk of the inner is no help" [ibid., p. 22]. According to McDowell, what's moving Wittgenstein's interlocutor is not a commitment to this or that picture of how one *knows* one's own mind: "The interlocutor's motivation here is not epistemological. A better word for it would be 'metaphysical'" [ibid., p. 20]. And McDowell thinks I go wrong in E&I — both on Wittgenstein's behalf and on my own — by explaining first-person authority in terms of expression, where I ought to be attending to the metaphysical issue that is the real source of the interlocutor's trouble at PI §304.

Now, I agree with McDowell about PI §304 insofar as I believe that Wittgenstein's interlocutor is not, there, worried *only* about epistemology. But it seems to me that McDowell misses — or, anyway, underplays — the extent to which this interlocutor thinks that what one says about psychological self-ascriptions is tied up with the metaphysics of the inner. (Moreover, I don't think *this* is something that Wittgenstein himself means to correct his interlocutor about.¹) If one looks at the passages that surround §304, it's hard to deny that Wittgenstein's interlocutor is thinking about — among other things, at least — the epistemology of self-awareness. Thus, in PI §303 he says, "I can only *believe* that someone else is in pain, but I *know* it if I am." And PI §305 begins this way:

"But you surely cannot deny that, for example, in remembering, an inner process takes place." — What gives the impression that we want to deny anything? When one says "Still, an inner process does take place here" — one wants to go on: "After all, you *see* it."

My point is *not* that, contra McDowell, "epistemological" is, after all, a better word for the interlocutor's motivation than "metaphysical." Rather, the interlocutor hears Wittgenstein's remarks about what might be described as the epistemology of psychological self-ascription as having disturbing metaphysical implications — as committing Wittgenstein to an untenable behavioristic denial of the inner. This doesn't happen only around PI §§303-305 (though it does happen there); it goes back to §244, where Wittgenstein introduces the idea that when a child learns to say that he's in pain, he is learning a new way to express his pain:

A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain-behaviour.

"So you are saying that the word 'pain' really means crying?" — On the contrary: the verbal expression of pain replaces crying and does not describe it.

When the interlocutor says what he does here, he is accusing Wittgenstein of embracing behaviorism; if the word "pain" meant crying, then it wouldn't refer to anything genuinely inner. This remark is of a piece with §307's "Are you not really a behaviourist in disguise?" His interlocutor sees that Wittgenstein is rejecting (what I call) *detectivism* — rejecting the very idea that when one avows a sensation, one is reporting a fact that one has, in some way or other, *learned*. (PI §246: "Other people cannot be said to learn of my sensations *only* from behaviour, — for I cannot be said to learn of them. I *have* them.") He (the interlocutor) can understand this only as a behavioristic denial of the inner, and this is precisely *because* he fails to understand (he doesn't have room in his picture of mind and behavior for) expression, i.e., for the way that

the inner can be manifestly present in what we do and say.² McDowell is right when he suggests (in the final sentence of his essay) that merely “insisting that self-reports of inner episodes express what they report” isn’t liable to do much good with Wittgenstein’s interlocutor — but, again, this is because the interlocutor doesn’t know *how* to think about expression. More than insistence *is* called for, but what’s needed is not to shift the conversation away from the topic of expression (and the expressive character of psychological self-ascriptions).

I think that part of what goes wrong in McDowell’s complaint that E&I is too focused on epistemology and the grammar of expression has to do with the way in which he understands the metaphysical fantasy that Wittgenstein’s interlocutor is gripped by in the vicinity of PI §304. “The reason we need philosophy in this area,” McDowell says, is “that we can easily be induced to mishandle the picture of the inner world” [this issue, p. 21], and we do this by “thinking we can acknowledge that an inner world is someone’s even while we insist that its inhabitants are present in it independently of its owner’s linguistic and conceptual competence” [ibid.]. McDowell’s focus is entirely on the relation between items in a stream of consciousness (e.g., sensations), on the one hand, and language/concepts, on the other. What goes missing here is that, for Wittgenstein, this way of mishandling the inner is a specific version of a broader metaphysical target, viz., the idea that sensations (and other mental states and goings-on) come, as it were, prior to and independently of *expressive* capacities, not only in creatures that can express their states of mind in language, but in animals and young children as well. If we fail to consider expression *as such* and focus only on the relation between sensations and linguistic/conceptual capacities, we’ll fail to understand why remarks like the following are to be found in this region of the *Investigations*:

Look at a stone and imagine it having sensations. — One says to oneself: How could one so much as get the idea of ascribing a *sensation* to a *thing*? One might as well ascribe it to a number! — And now look at a wriggling fly and at once these difficulties vanish and pain seems able to get a foothold here, where before everything was, so to speak, too smooth for it.

And so, too, a corpse seems to us quite inaccessible to pain [PI §284].

The difference between a stone and a fly is not that the fly has linguistic/conceptual capacities. But a fly’s wriggling can be seen (maybe) as a kind of expression. (Relatedly, what the passage calls attention to in connection with corpses is not that they are unable to talk, but that they cannot express sensations at all — not even by wriggling.)³

To sum up, I believe that McDowell is right to insist that “[t]he reason we need philosophy in this area” is “not specifically epistemological” [this issue, p. 21]. But he is wrong to think that E&I’s emphasis on the expressive

character of psychological self-ascription amounts to a slighting of what's at stake metaphysically in and around PI §304. Arriving at a clearer view of the expressive dimension of psychological self-ascriptions is not a "specifically epistemological" achievement.

II. REPLY TO BARRY STROUD

Barry Stroud asks many questions about E&I in his fascinating and sympathetic essay — too many questions for me to address in a short reply such as this. Still, it would be reasonable for you to expect that I'd try to answer a few of them, at least. I'm afraid that's not going to happen. Instead, I shall call your attention to two moments in his essay when Stroud seems either to be agreeably summarizing my view or, anyway, stating something that he and I would agree about. In each case, I can imagine a reading of what he says according to which this *is* what he's doing. But, in each case, what seems to me the more natural reading is one according to which he is, in fact, disagreeing with something that is essential to my position. When I was thinking about how best to reply to Stroud's essay, it struck me that by discussing these two moments in it, I might elucidate some of what is difficult to understand in E&I. So, with that as my aim, here goes.

Early in his essay, Stroud describes a child who has learned: (1) to say what colors objects have and (2) to ascribe attitudes both to herself and others. Once she can do these things, she can self-ascribe knowledge of an object's color "as soon as she has determined what colour that object is," without any "further investigation" [this issue, p. 26]. The child merely needs to see, e.g., that the toy in front of her is red, and this suffices for her to know that she knows or believes that that toy is red. Stroud goes on:

In this simple case, she has only to look at an object and see and thereby know what colour it is. With more complicated questions it is not that easy. But even when complex thought and elaborate investigation are called for, a person's coming to know or believe what she does is something she is responsible for in the sense that it is the outcome of her own active reflection and deliberation. That is what puts her in a position to ascribe that knowledge or belief to herself [ibid., p. 26].

Stroud doesn't indicate that, here, he means to be taking issue with, or questioning, the position I defend in E&I. And I do agree that once someone sees, or otherwise makes up her mind, about the color of an object, she needn't do anything *more* in order to know — or anyway, be in a position to self-ascribe — what she believes about its color.⁴ But Stroud says of the child he is describing that "what puts her in a position" to self-ascribe a belief is the fact that

her coming to believe what she does is “the outcome of her own active reflection and deliberation.” And this strikes me as misleading — at least given what I claim about psychological self-ascriptions.

Richard Moran is someone who, it seems to me, might be well-described as thinking that rational deliberation is what puts one in a position to first-personally self-ascribe beliefs and other attitudes. On his view, one enjoys “genuine first-person awareness” [Moran (2001), p. 107] of, e.g., a belief or a fear only when one is able to avow it by deliberately addressing the question of what to believe or to fear. But I hold that regardless of whether someone is avowing a deliberated belief, an irrational fear, or a headache, she speaks with an authority that lies in (we might say) her avowal’s expressing the very state that it’s an avowal *of*. Deliberation might put someone in a position to self-ascribe a rational attitude in the sense that listening to loud music might put someone in a position to self-ascribe a headache.⁵ But one is (ordinarily) in a position to speak with first-person authority about one’s states of mind regardless of whether they are the results of — or even consistent with — active reflection or deliberation.⁶ On my view, this authority represents a unitary phenomenon with a single explanation.

* * *

Stroud writes:

I think someone who says ‘I have got a splitting headache’ is certainly expressing her pain. But she does so not by holding her head or moaning, but by uttering an articulate sentence whose structure and meaning she understands and which is either true or false. She asserts that sentence, or puts it forward as true. That is what she learned to do when she learned to ascribe mental states to herself rather than simply expressing them non-linguistically. Her saying that she has a splitting headache is not to be understood as a “mere effect” of her being in that state of mind or feeling, as a wince or a moan might be [this issue, p. 30].

Here again, Stroud doesn’t indicate that he means to be disagreeing with me.⁷ And of course, I wholeheartedly agree with most of what is said in this passage. But Stroud here suggests that when someone with a headache winces or moans in pain, the wince or the moan might be a “mere effect” of the headache. This is not something that I would want to say. In §5.3 of E&I, I argue that a mental state and its expression — regardless of whether the latter takes linguistic form — may be understood as making sense together, in light of each other, in something like the way in which the two parts of a single sentence make sense together. While I have no objection to also saying that headaches *cause* moans, the relationship between them is not, on my view, *merely* causal (anymore than the relationship between, e.g., wanting something

to eat and making oneself a sandwich is merely causal⁸). Moaning — even a dog’s moaning — is not well-described as a “mere effect” of the pain that it expresses.⁹

In rejecting the idea that the relationship between a mental state and its expression — whether linguistic or not — is merely one of cause and effect, I am (again) following Wittgenstein. Consider *Zettel* §53, which begins as follows:

I expect an explosion any moment. I can’t give my full attention to anything else; I look in a book, but without reading. Asked why I seem distracted or tense, I say I am expecting the explosion any moment.

Here, someone first expresses an expectation in non-verbal behavior; then he avows it. We can think of what immediately follows (still in Z §53) as comprising two competing suggestions about how to understand the ways in which expectation, non-verbal behavior, and avowal are related to one another:

— Now how was it: did this sentence describe that behaviour? But then how is the process of expecting the explosion distinguished from the process of expecting some quite different event, e.g. a particular signal? And how is the expectation of *one* signal distinguished from the expectation of a slightly different one? Or was my behaviour only a side-effect [*Nebenerscheinung*] of the real expectation, and was that a special mental process?

The first suggestion made is that the sentence, “I am expecting the explosion any moment,” is just a description of the speaker’s restless behavior. (This is the same sort of behavioristic suggestion that Wittgenstein’s interlocutor worries he might be making in connection with pain at PI §244.¹⁰) The second suggestion is that the non-verbal behavior is just a “side-effect” — merely the causal upshot — of the expectation itself. Wittgenstein takes these to be two horns of a *false* dilemma. (By grasping the second horn, Wittgenstein’s interlocutor winds up committed to a problematic sort of platonism.) In my reply to John McDowell’s contribution to this symposium, I said that Wittgenstein’s interlocutor has no room in his picture of mind and behavior for expression — and so no room for Wittgenstein’s oft-repeated suggestion that psychological self-ascriptions are, or are akin to, expressions. The interlocutor has no room for expression *because* he is in the grip of this false dilemma — because he cannot see any alternative to behaviorism other than the idea that a mental state and its expression are related *merely* as cause and effect.

III. REPLY TO JESÚS VEGA-ENCABO

At the end of Chapter 6, I discuss the case of Hank, who learns that a sharp pain behind the knee is a telling symptom of Green’s disease. Hank now

experiences such a pain, and he reasonably infers that he's liable to have the disease. I consider several questions about Hank, one of which is whether or not he should be said to *know* that he has a pain in the back of his knee. Upon raising this question, I ask the reader to imagine three philosophers. All of them accept what I claim in the bulk of the book about psychological self-ascriptions and first-person authority. But they disagree about how the word "know" ought to be used. Philosopher 1 says: "It's wrong to speak of Hank as either 'knowing' or 'not knowing' that he's in pain. Such talk is, at best, unnatural and misleading and, at worst, sheer nonsense." Philosophers 2 and 3 are less concerned about lapsing into nonsense (perhaps they aren't fans of Wittgenstein). Philosopher 2 says that — because Hank's self-ascription is not based on epistemic grounds — "he doesn't know that he's in pain." Philosopher 3 says: "The mistake is to think that knowledge always requires epistemic justification. Hank does know that he's in pain." In the final paragraph of Chapter 6, I write, "I don't see that these philosophers need to be disagreeing about anything of genuine philosophical import" [E&I, p. 152].¹¹

Jesús Vega-Encabo believes that something of philosophical import *is* at stake. Moreover, a central aim of his essay is to argue against Philosophers 1 and 2 and in favor of 3. He means to show that "taking the assertoric dimension of avowals seriously means committing oneself to an epistemic description of first-person self-ascriptions" [this issue, p. 46]. Now, this isn't a conclusion that threatens any claim to which I'm attached, and I'd be happy to be convinced that Philosopher 3's position is superior to that of Philosophers 1 and 2. But I don't find myself convinced by the reasons Vega-Encabo adduces in support of this conclusion. For example, he points out that "[c]haracteristically, assertions are moves in an epistemic game" [ibid., p. 40]. And he writes:

One of the basic functions of assertions is to transmit possible knowledge. So if self-ascriptions are also assertions, nothing is more natural than to consider them as manifestations of knowledge [ibid., p. 42].

But neither of our first two philosophers needs to deny that assertions are *characteristically* manifestations of knowledge. The question is what to say about a special class of assertions — assertions like "I'm in pain" — which (all three philosophers agree) have an uncharacteristic epistemic profile *at least* insofar as they are not based on anything like observation, inference, or testimony.

In what remains of this reply, I'll focus on an argument in favor of Philosopher 3 that Vega-Encabo offers in §4 of his paper. Here is a paraphrase of the argument: Imagine that, at time *t*1, I learn, on the basis of solid evidence, that I'm unconsciously angry at my brother. Imagine, moreover, that this realization isn't enough to make my anger conscious; my situation at *t*1 is that I know something about an unconscious mental state of

mine. Years go by, during which I work through my feelings toward my brother, and by t_2 , my anger has come to consciousness; I'm now able to speak about it with first-person authority. So at t_1 , I *knew* that I was angry. What is my epistemic situation at t_2 ? Vega-Encabo states the conclusion of his argument in the form of a question: "Could I possibly be taken as not knowing the proposition that I previously knew?" [this issue, p. 44].

What should we say about this argument? The first thing to notice is that there is an ambiguity in Vega-Encabo's formulation of the question with which he concludes the argument. He means to suggest that Philosophers 1 and 2 are committed to claiming, implausibly, that, at t_2 , I should "be taken as not knowing the proposition that I previously knew." But what *is* "the proposition that I previously knew"? Barring memory lapses, at t_2 , I still know that, at t_1 , I *was* (unconsciously) angry at my brother; neither Philosopher 1 nor Philosopher 2 has any reason to deny this. In order to avoid this way of hearing his question, Vega-Encabo might restate it as follows: "Could I be said not to know that I am angry at t_2 , given that, at t_1 , I knew that I was angry?" Now, it doesn't seem to me that Philosopher 1, anyway, should be disturbed by this question. Indeed, I imagine Philosopher 1's replying to it as follows:

That's *not* what I would say about you at t_2 . I hold that it doesn't *make sense* to say of you at t_2 , when your anger is conscious, "He doesn't know that he's angry." When Wittgenstein writes, "It can't be said of me at all (except perhaps as a joke) that I *know* I am in pain" [PI §246], he doesn't mean that it *can* be said of me that I *don't* know I'm in pain.

What about Philosopher 2? Here I think Vega-Encabo's argument is more helpful. It would surely be *misleading*, anyway, to say of someone whose anger was entirely conscious, "He doesn't know that he's angry" — for this would suggest precisely that his anger was *unconscious*. Vega-Encabo's discussion helps to bring out that Philosopher 2 is using the word "know" in a way that is far removed from how we ordinarily use it. Maybe this suggests that Philosopher 2 is confused in a serious way. But, whatever we say about Philosopher 2, I don't think that Vega-Encabo makes a convincing case (either with this argument or elsewhere in his paper) for thinking that Philosopher 1 has got things wrong. So I don't think his paper shows that "taking the assertoric dimension of avowals seriously means committing oneself to an epistemic description of first-person self-ascriptions" [ibid., p. 46].

IV. REPLY TO JOSEP CORBÍ

In the introduction to E&I, I describe the case of Max, who has mistakenly purchased tickets for two different concerts that will take place at the

same time on the same evening. When the question arises about which of the two events he plans to attend, Max says, “I intend to see Dylan.” Max’s case, as I describe it in E&I, exhibits two features that, throughout the book, I take to be characteristic of “first-person authority”: (1) If someone wants to know which concert Max intends to attend, Max himself is the best person to ask, and (2) Max is able to say what he intends without relying on the sort of behavioral evidence that another person would need in order to responsibly ascribe an intention to him.

After reviewing this example, Josep Corbí writes, “Finkelstein commits himself to quite a substantial thesis, namely: what is true about intentions concerning which concert to attend, also *applies unrestrictedly to hopes, fears, desires, beliefs, moods, emotions, sensations, and passing thoughts*” [this issue, p. 52]. On one reading of this sentence, it is correct. I *am* committed to the thesis that *insofar* as a particular self-ascription of hope, fear, or desire (etc.) exhibits the two features of “first-person authority” mentioned above in connection with Max’s case, this self-ascription — like Max’s self-ascription of intention — may be understood as an expression of the very state (or event) that is being self-ascribed. I’m committed moreover, to the idea that once we get this expressive side of psychological self-ascriptions into focus, we shouldn’t find “first-person authority” (again, understood in terms of those two features) philosophically problematic. But Corbí’s characterization of the thesis to which I’m committed is susceptible to other readings. A thesis to which I’m *not* committed might be put as follows: *every* distinctively first-personal feature of a subject’s relation to her own state of mind may be understood by attending to the expressive dimension of psychological self-ascriptions and is accounted for in E&I. I’ll return to this point.

Corbí distinguishes between two kinds of (what *he* calls) “first-person authority” — an “epistemic” kind and a “practical” kind [ibid., p. 53]. If our focus is *epistemic* authority, then “evidence is at issue, even if it is from the viewpoint of its irrelevance” [ibid.], whereas with *practical* authority, “the capacity to shape one’s own psychological and behavioral dispositions is the central concern” [ibid.]. Now, I have no objection to someone’s using the phrase “first-person authority” differently from the way that I do — to refer to two separate first-personal phenomena. Given Corbí’s terminology, one could say that E&I provides an account of epistemic first-person authority, but not of practical first-person authority. This is fine with me; hence the conclusion of the preceding paragraph. But, in his essay, Corbí claims that I equivocate between these two kinds of first-person authority, and I don’t see that this is so.

Why does Corbí think I equivocate? He understands the distinction between epistemic and practical first-person authority to be connected to another distinction that he draws in his paper — between what he calls “trivial” cases of first-person authority (Max’s case is, according to Corbí, a trivial one) and

“more complex ones” [this issue, p. 53]. He suggests that in E&I, I equivocate by “assimilat[ing] the sort of first-person authority that trivial cases *display*” [ibid., pp. 55] with a kind of authority that is at issue only in complex cases. But: he defends this claim in large part via a discussion of a case that I don’t consider at all in E&I (or elsewhere) — the case of “Frau Bruggen.” In effect, Corbí imagines what I *would* say about Frau Bruggen and then complains that I would mistake the “practical” authority that Frau Bruggen lacks for the kind of “epistemic” authority that I focus on in E&I. It’s hard for me to imagine how this critical strategy could be carried out persuasively; in any case, I don’t think Corbí manages it.

Setting aside the question of whether the case of Frau Bruggen reveals (or could reveal) an equivocation on my part, it seems to me that some of what Corbí says about the case is, in a significant way, off-key. As Freud tells the story, Frau Bruggen is unconsciously in love with her son-in-law. At the start of §4 of his paper, Corbí notes that she might become aware of her unconscious desire for her son-in-law without the desire’s becoming conscious. So far, so good. But Corbí goes on to say:

[W]ith regard to this desire, we must also grant Frau Bruggen the same trivial ability that was ascribed to...Max. She would sincerely claim that it is a most monstrous desire and still her plight would be that she is unable to impose her authority upon such a desire, which far from extinguishing as a result of Frau Bruggen’s view about its monstrous character, keeps on pressing; its pressure being partly released through her obsessive behavior. So, it seems that a condition for Frau Bruggen’s predicament is (a) that there are some beliefs and desires with regard to which she is in possession of the trivial kind of first-person authority which...Max’s [case exemplifies], but (b) she is unable to impose her practical authority with regard to the psychological and behavioral dispositions that are relevant to such beliefs and desires [this issue, pp. 59-60].

According to Corbí, once she has become aware of her unconscious desire, Frau Bruggen would have “the same trivial ability that was ascribed to...Max.” But Max is able to express his state of mind — his intention to see Dylan — by self-ascribing it. Frau Bruggen might “sincerely claim that it is a most monstrous desire” that she has become aware of in herself, but if this desire remains unconscious, then her self-ascription of it *won’t* express it. It’s absolutely central to what my position implies about a case like Frau Bruggen’s that Corbí is wrong when he says that “with regard to this desire, we must also grant Frau Bruggen the same trivial ability that was ascribed to ... Max.” Of course, it doesn’t follow from this that Corbí is wrong to think that Frau Bruggen suffers from an absence of what he calls “practical first-person authority.” But the suggestion made in the passage I just quoted — that we must understand what goes missing here in terms of “practical first-person

authority” *because* Frau Bruggen already enjoys the sort of first-person authority exhibited by Max — is, on my view, mistaken.

V. REPLY TO JOSEP PRADES

In a paper I’ve long admired, Cora Diamond (1989) writes:

Terms like ‘logically possible’, ‘conceptually possible’, when they occur in philosophy, often indicate some kind of confusion. It is misleading to use them in giving Wittgenstein’s views, if we are concerned with passages in which he did not use them [Diamond (1989), p. 20].

The point that Diamond has in mind here — about terms like “logically impossible” — is one that Wittgenstein sometimes makes in connection with talk about what “cannot be imagined.” In the passage that follows, he is discussing a philosopher who says, “I cannot imagine having Smith’s pain”:

[I]t looks as if there is a certain goal that we could get nearer and nearer to, whereas as a matter of fact there is none. It is not that the man who said he could not imagine having Smith’s toothache was debarred from imagining anything. He described all the things he imagined and then rejected them, so that it seemed as if there was a goal out of reach. But it now appears that there was no such goal, but only a verbal expression “I have Smith’s pain”, which, by the way is perfectly good English (LPE).

It’s not very hard to imagine situations in which one might say, “I have Smith’s pain.” Here’s one: My friend Smith tells me that ever since her recent trip to South America, she’s been experiencing a peculiar, fluttery sort of sharp pain in the back of her knee. She says she’s worried that she might have contracted a disease there. She tells me this just *after* I’ve had a sip of the orange juice that she’s been drinking. Five days later, I experience a sharp, fluttery sort of pain in the back of my knee, and I say: “Uh oh. I have Smith’s pain.” Now, the philosopher who said, “I cannot imagine having Smith’s pain,” might respond to this story by saying: “That’s not what I meant by ‘having Smith’s pain’ when I said it was unimaginable. Of course I can imagine *that*.” But if we ask him what he *did* mean by “having Smith’s pain,” he’s put in an awkward position. For he’s inclined to reject any coherent meaning that one might give to those words as “not what I meant.” The right conclusion to draw about such a philosopher is *not* that he meant something by “having Smith’s pain” that has the peculiar feature of being “unimaginable” or “logically impossible.” It’s that he didn’t mean anything — didn’t settle on any particular thing *to* mean — by those words. He may have *thought* he meant something, but there are illusions of meaning. (And one of the aims of

Wittgenstein's philosophy — perhaps its central aim — is to help us see such illusions for what they are.)

In his essay, Josep Prades repeatedly stresses his agreement with me concerning how to understand Wittgenstein's remarks about rule following — how to understand what Wittgenstein meant by them and also what we should learn from them. But he thinks that there is a difficulty that I fail to deal with in E&I — an obstacle that stands in the way of my being entitled to the conclusions that I would draw from Wittgenstein's remarks. He puts the difficulty this way:

[F]or everything he [Finkelstein] says, nothing in the past practice seems to be able to protect us against the intelligibility of completely deviant ways of applying language to new situations The fact that we have been led by our nature and the processes of socialization to understand some linguistic expressions in a certain way without the need of interpretation, does not seem to entail that we cannot conceive of people who, for instance, tomorrow would start reacting in a completely different, bizarre, and deviant way [this issue p. 68].

Prades thinks that in order to meet this difficulty, it must be shown that there is a limit to how much of this sort of deviance is possible.¹² Even if we can imagine systematically bizarre, deviant reactions to *linguistic* expressions of *semantic* content, he thinks we must be brought to recognize that the same is not true about *non-linguistic* expressions of *intentional* content.¹³ He writes:

There is something that cannot be the content of our imagination: some animal that systematically perceives intentionality in a completely bizarre way. There cannot be such a thing, and there cannot be any coherent conception of it. Simply, in basic cases of intentionality, there cannot be alternative, systematically bizarre perceptions of content that would count as perception of content... [T]here is no possible world in which intentionality is expressed in completely different ways to the ways we consider normal [this issue, p. 72].

In light of what I said earlier about Diamond and Wittgenstein, what Prades says here should, I think, seem unsatisfactory on the face of it. He says that there is "something that cannot be the content of our imagination" — something that, even so, he doesn't seem to have trouble *describing*. Indeed he prefaces one description of it with the word "Simply." Yet, he also says that "there cannot be any coherent conception" of this something. In trying to solve the problem that he sets out, Prades gets himself into just the sort of trouble that Wittgenstein and Diamond have in mind in the passages that I quoted earlier.

But even if his solution is flawed, Prades's problem might still be something that we need to worry about. How should we respond to it? Well, we might ask: Why should we feel threatened by skepticism if we can "conceive of people who, for instance, tomorrow would start reacting in a completely

different, bizarre, and deviant way”? Why not just say that these imagined people would have lost the ability (if they ever had it) to *understand* what we were saying — and perhaps what our non-linguistic expressions meant as well? (I once worked in a hospital with groups of aphasics; they never threatened my confidence that I meant something by my words. Should they have?) Prades himself considers a response somewhat like this. The passage from p. 68 of his paper that I quoted in the preceding paragraph continues as follows:

Of course, we can always say that those strange, deviant people would be giving different meanings to our words. This might be true, but the problem would still be that those alternative meanings would be equally rooted...in the common past practice of using language [ibid.].

Here, I think we come to the crucial movement in this conjuring trick. Why does Prades suppose that the strange, alternative meanings assigned to our words by bizarre, deviant people would be just as rooted in our past practice of using language as the meanings *we* take our words to have? (Should I have thought that my aphasics’ responses to linguistic stimuli were just as rooted in our common past practice of using language as my own?) If we conceive of our past “practice” of using language in entirely *non-semantic* terms — so, as a history of generating “sounds, ink-marks” [PI §431] devoid of any significance — then bizarre (perhaps quus-like) interpretations of our words seem to be just as rooted in our “practice” as anything else. But if we allow that a description of linguistic practice might look like this: “English, French, and German speakers use ‘plus’ to mean *plus*,” then bizarre, deviant interpretations are *not* “equally rooted...in the common past practice.” This is to say: the issue that Prades raises — that, for all Finkelstein says, we seem able to “conceive of people who, for instance, tomorrow would start reacting in a completely different, bizarre, and deviant way” — presents a *problem* only *given* the interlocutor’s picture in PI §431, the picture of an ever-present gulf between dead signs and appropriate reactions to them. I believe that the right response to this supposed problem is *not* to show that “[t]here is something that cannot be the content of our imagination: some animal that systematically perceives intentionality in a completely bizarre way.” It is, rather, to undermine the picture that would lead someone to say that systematically bizarre interpretations of our words would be “equally rooted ... in past practice.” I discuss the ways in which Wittgenstein does this in §4.1 of E&I.¹⁴

Department of Philosophy
The University of Chicago
 1115 E. 58th St. Chicago, IL 60637, USA
 E-mail: davidf@uchicago.edu

NOTES

¹ Consider, e.g., the following from LWPPII, p. 64:

“In investigating the laws of evidence for the mental, I am investigating the essence of the mental.” Is that true?

Yes.

² In E&I, I try to bring out how the interlocutor’s inability to get a grip on the very idea of expression is of a piece with his inability to see “that there is a way of grasping a rule which is *not an interpretation*” [PI §201]. It’s not an accident that Wittgenstein starts talking about the expressive character of psychological self-ascriptions soon after an extended discussion of rule-following in the *Investigations*. In both discussions, the epistemological issues are tightly bound up with the metaphysical ones.

³ The point I’m making (or gesturing toward) in this paragraph is connected — in ways that I can’t spell out here — to my introducing talk about a “logical space of animate life” in §5.6 of E&I.

⁴ I’m going to bracket the difficult question of whether we enjoy first-person authority about our own states of knowledge. (This is one of several questions raised by Stroud’s piece that deserves its own essay.)

⁵ This is not to deny that there are, as it were, *dependence* relations between first-person authority and rational deliberation that are worth talking about. (I hope to say more about these in future work.)

⁶ For more on this point and its implications, see E&I’s postscript (or the summary of it in the précis that appears at the start of this symposium). For more still, see Finkelstein (forthcoming).

⁷ He *is* working his way up to questioning something that I say about someone who complains of a splitting headache. But in the quoted passage, he doesn’t register that he’s reached the point of disagreeing with me, or questioning me, *yet*.

⁸ On my view, this sort of rational relation is a kind of expressive relation. One way in which I can express my desire for something to eat is by making myself a sandwich.

⁹ I *can* imagine a reading of the final sentence of the passage I quoted from Stroud that is innocent from my perspective, one according to which he means nothing more — when he says that a wince might be the “mere effect” of a headache — than that wincing are *not* assertions. But on what seems to me the most natural reading of that sentence, it’s crucial to my position that what he suggests about the relation between pains and wincing is false.

¹⁰ As I mentioned in my reply to McDowell, Wittgenstein’s interlocutor is continually worrying that he (Wittgenstein) is lapsing into a behavioristic denial of the inner. In Z §53, the behavioristic suggestion is, in effect, met with an objection — one that might be spelled out as follows: The non-verbal behavior exhibited in this case (picking up a book and then immediately putting it down, etc.) seems to be consistent with all sorts of expectations. It doesn’t look as if this behavior is, as it were, sufficiently *fine-grained* to license any *particular* ascription of expectation. Yet the subject

says that he expects a very particular thing — an explosion at any moment. And we're not inclined to doubt his word.

¹¹ Vega-Encabo (mis)characterizes me as saying something stronger than this — as claiming that these three philosophers “do not disagree ‘about anything of genuine philosophical import’” [this issue, p. 36].

¹² “In order to preserve a full-blooded notion of determinate meaning, we must show that certain abilities that are grounded on contingent facts of human nature set the limits on what is conceptually possible” [this issue, pp. 67-68].

¹³ Thus there is, he says, an “asymmetry between intentional content and linguistic meaning” [this issue, p. 71].

¹⁴ I'm grateful to Stina Bäckström, Irad Kimhi, and Thomas Lockhart for helpful comments on drafts of these replies.

REFERENCES

Abbreviated References to Wittgenstein's Writings:

- LWPPII: WITTGENSTEIN, L. *Last Writings on the Philosophy of Psychology*, vol. 2
C.G. Luckhardt and M.A.E. Aue (trans), G.H. von Wright and H. Nyman
(eds.), Oxford: Blackwell, 1992.
- PI: *Philosophical Investigations*, G.E.M. Anscombe (trans.), New York: Macmillan Publishing Co., Inc., 1958.
- LPE: *Lectures on Personal Experience, Cambridge, Michaelmas Term, 1935*,
James Conant and Cora Diamond (eds.), Cambridge, MA Harvard University Press, forthcoming.
- Z: *Zettel*, G.E.M. Anscombe (trans.), G.E.M. Anscombe and G.H. von Wright (eds.), Berkeley: University of California Press, 1981.

Other References:

- DIAMOND, C. (1989), “Rules: Looking in the Right Place,” in *Wittgenstein: Attention to Particulars*, Phillips, D.Z. and Winch, P. (eds.), New York: St. Martin's Press, pp. 12-34.
- FINKELSTEIN, D.H. (2003), *Expression and the Inner*, Cambridge, MA: Harvard University Press.
- (forthcoming), “From Transparency to Expressivism,” in *Rethinking Epistemology*, Guenter, A. and Conant, J. (eds.), Berlin: De Gruyter.
- MCDOWELL, J. (1996), *Mind and World*, Cambridge, MA: Harvard University Press.
- (2009), “Avoiding the Myth of the Given,” in *Having the World in View*, Cambridge, MA: Harvard University Press, pp. 256-272.
- MORAN, R. (2001), *Authority and Estrangement: An Essay on Self-Knowledge*, Princeton, NJ: Princeton University Press.
- SELLARS, W. (1997), *Empiricism and Philosophy of Mind*, Cambridge, MA: Harvard University Press.