

and so on. But if I am convinced that a mouse cannot come into being from these things, then this investigation will perhaps be superfluous" (§52). Here, the mouse would be the meaning, and the use, the grey rags and dust. For an illuminating discussion of the significance of this passage, see Diamond (1992: 39–73).

13. The expression is from Stanley Cavell.
14. You might say that such pupils represent "hyperbolic" possibilities of misunderstanding. This brings to mind Descartes' "hyperbolic doubt". It would be useful to compare the distinction between ordinary and hyperbolic possibilities here with the distinction between ordinary and hyperbolic epistemic possibilities, such as the possibility that one is dreaming as it functions in Cartesian scepticism about what we can know.
15. "But do you really explain to the other person what you yourself understand? Don't you get him to *guess* the essential thing? You give him examples, – but he has to guess their drift, to guess your intention? – Every explanation which I can give myself I give to him too" (§210).
16. "But if a person has not yet got the *concepts*, I shall teach him to use the words by means of *examples* and by *practice*. – And when I do this I do not communicate less to him than I know myself" (§208).
17. "Philosophy simply puts everything before us, and neither explains nor deduces anything. – Since everything lies open to view there is nothing to explain. For what is hidden, for example, is of no interest to us" (§126).
18. One thing that is shown by what I am calling "the example of the wayward pupil" is that initiation of a child into the practice of our language depends on the child's ability, and willingness, to catch on to the practices through a certain training. And this depends in turn on the child's sharing with us certain natural responses, or a sense of what is and is not natural. The child must find it natural to look in the direction in which our finger points, agree with us in our judgements about what counts as going on in the same way, find salient in a situation what we find salient there, be comforted by what we are comforted by, feel pain by what gives us pain, laugh at what we laugh at, and so on. For a bravura treatment of this point, see Cavell (1979b: 168–90).
19. "In giving explanations I already have to use language full-blown (not some sort of preparatory, provisional one); this by itself shews that I can adduce only exterior facts about language" (PI §120).
20. "One might say: an explanation serves to remove or to avert a misunderstanding – one, that is, that would occur but for the explanation; not every one that I can imagine ... The sign-post is in order – if, under normal circumstances, it fulfills its purpose" (PI §87).
21. See Cavell (2005: esp. 112–14).
22. In addition to the writings by Cavell, Diamond, Goldfarb and Rhees mentioned here, I have found helpful guidance from the work of Barry Stroud and John McDowell. See, in particular, Stroud (2000) and McDowell (2009). I am grateful for discussions with Kelly Jolley on the issues discussed in this chapter.

Expression and avowal

David H. Finkelstein

SIXTEEN

But that which is in him, how can I see it? Between his experience and me there is always the expression!
 Here is the picture: He sees it immediately, I only mediately.
 But that's not the way it is. He doesn't see something and describe it to us.
 (LWPP II 92)

I

I am often able to say what is on, or in, my own mind, that is, what I want, believe, fear, expect, intend or hope; whether I am feeling joy or pain; whether I like the taste of this wine or find that joke funny. I seem, moreover, to manage this sort of self-ascription, or avowal, without needing to rely on the evidence that other people require in order to ascribe mental states to me. In his late work, Wittgenstein often writes about psychological self-ascriptions. Again and again he suggests that, in doing philosophy, we are liable to cling to one or another bad explanation or misleading picture of them. He aims to reorient our thinking about avowals by urging us to view them as expressions, and so to see a sincere utterance of "I am in pain" as akin to a pained wince or groan. My aim in what follows is to provide an introduction to this strand in his writing.

A good place to begin is with the quotation above – our epigraph – which was written during the final two weeks of Wittgenstein's life. In it, he sketches a "picture" that we can think of as comprising two claims. The first is that a person knows what is going on in his own mind

(“that which is in him”) by a kind of inwardly directed observation or perception. The second claim is that one cannot know what is going on in the mind of another person in as direct a manner as this. So: while you are able to perceive your own mental goings-on, my access to your psychology is always mediated by your behaviour. You “see” your own, for example, pains; I see only the behaviour to which your pains give rise. In this sense, your behavioural expressions of pain stand *between* me and that which they express. (“Between his experience and me there is always the expression!”)

Of course, the idea is not that you literally *see* your own states of mind. (You do not use your eyes.) It is, rather, that you have a kind of access to them that is, in significant respects, akin to your visual access to what is before your eyes. Philosophers and psychologists have held a range of positions that fit this characterization. According to some – for example Bertrand Russell (1912) – while “inner sense”² is in certain notable respects like the outer senses, it is nonetheless more direct and less fallible than seeing, hearing, smelling and so on ever are, or could be.¹ According to other (typically more recent and more naturalistic) theorists (see e.g. Armstrong 1968; Humphrey 1986; Lycan 1996), inner sense is just one fallible perceptual modality among others – a process that is epistemically on all fours with seeing or hearing, except that it happens to be directed inward, towards goings-on in the mind/brain rather than outward towards the external world. As we shall see, Wittgenstein takes the whole gamut of such positions to be confused. Indeed, the whole gamut of such positions is a species of a broader genus that he would have us reject. Thus he writes: “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” (PI §246).² Wittgenstein rejects not only views according to which I learn of my own (e.g.) sensations by some sort of inner sense, but the very idea that I learn of them *at all*.

But now, this might seem puzzling (or worse). After all, Wittgenstein would not deny that I am generally able to say when I am in pain and when I am not. Does my having this ability not entail that, somehow or other, I learn of my own sensations?

II

We can begin to address this question, along with our overall topic, by considering the following, from §244 of *Philosophical Investigations*:

[W]ords are connected with the primitive, the natural, expressions of the sensations and used in their place. 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.

Think of this passage as suggesting a way to understand (what Wittgenstein would call) the grammar of pain self-ascriptions. What a child learns when he learns to avow his own pains is, so the thought goes, “new pain-behaviour”, that is, new ways to express pains. Whereas, before, the child might express his pains by wincing or crying out, now he can also express them by saying, for example, “I bumped my toe, and it hurts!” The child learns to express his pains *by* self-ascribing them.

While *Philosophical Investigations* §244 concerns the grammar of sensations in particular, Wittgenstein elsewhere indicates that the expressive character of their deployment in the first person is a mark of psychological attributions more generally. For example, he writes:

The statement “I am expecting a bang at any moment” is an *expression* of expectation. (Z §53)

For even when I myself say “I was a little irritated about him” – how do I know how to apply these words so precisely? Is it really so clear? Well, they are simply an expression. (LWPP II 70)³

Plan for the treatment of psychological concepts.

Psychological verbs characterized by the fact that the third person of the present is to be verified by observation, the first person not.

Sentences in the third person of the present: information. In the first person present: expression. ((Not quite right.))

The first person of the present akin to an expression. (Z §472)

I shall leave for later *Zettel* §472’s doubly parenthetical “Not quite right.”⁴ For now, it suffices to say that, according to Wittgenstein, we should understand a very wide range of psychological self-ascriptions – not only ascriptions of sensation, but of attitude and emotion – as akin to pained winces, desirous glances and angry door-slamming.

Clearly, what such a suggestion amounts to depends a good deal on how one chooses to address the following pair of questions: (i) in precisely which respects are psychological self-ascriptions akin to wincing, smiles and door-slamming? And (ii) how should we think about wincing, smiles, door-slamming and so on? What does it mean to call such behaviours *expressions*? In what remains of the present section, I shall say a little about Wittgenstein on question (ii). Later (in §1V), I shall argue against a widespread understanding of him that, I believe, goes wrong by misrepresenting his answer to question (i).

What does it mean to call something that a person does with, say, his eyes and mouth an expression? Wittgenstein writes:

“We see emotion.” – As opposed to what? – We do not see facial contortions and make inferences from them (like a doctor framing a diagnosis) to joy, grief, boredom. We describe a face immediately as sad, radiant, bored, even when we are unable to give any other description to the features. – Grief, one would like to say, is personified in the face. (Z §225)

We could say that for Wittgenstein, an expression is, typically, the perceptible emergence of some psychological phenomenon (or phenomena) in a creature's doings. A facial expression can make someone's state of mind immediately manifest to others. This does not, I hope, sound like a technical or idiosyncratic notion of expression. Pre-philosophically, we take it for granted that it is possible – and not at all uncommon – to see in a face, or hear in a voice, that someone else is, for example, pleased, anxious, angry, in pain or dubious about what we have just said. When we are doing philosophy, however, we are liable to conclude that this cannot really be the case – that whenever we take ourselves to see psychology in someone's face, we are in fact seeing no more than physical movements (“facial contortions”) and making inferences about the hidden states and processes that might have caused them. Wittgenstein's criticisms of the ways in which we arrive at this conclusion run deep in his work, and it goes beyond the scope of the present chapter to explicate them. Any such explication would require discussion of his much-debated remarks about signs, meaning and rule-following, for he thinks that we are inclined to just the same kinds of confusion whether we are talking about the relation between signs and their semantic significance or that between behaviour and its psychological significance. For now, it will have to suffice for me to note that, according to Wittgenstein, just as we can (if we are rightly situated, know the language, etc.) see meaning in the words on a page, we can

(if we have the right sensibilities, know enough about the person, etc.) see psychology expressed in someone's behaviour.

Wittgenstein's interlocutor tends to take it for granted that all we see when we look at a face are what amount to contortions – movements that are, in themselves, devoid of psychological content. For this reason, he suspects that Wittgenstein's willingness to allow that we can see psychology in a face commits him to a kind of reductive behaviourism. This interlocutor, in effect, reasons as follows: “If Wittgenstein imagines that we can see psychology in mere movements, he must hold that psychological states and processes are nothing more than behaviour; he must be ‘a behaviourist in disguise.’”⁵ From the perspective of Wittgenstein's interlocutor, we must choose between an implausible reductive behaviourism and a position according to which we can, at best, make inferences about hidden states and processes that cause the in-themselves-psychologically-neutral movements and noises that we perceive when we look at and listen to other people. In order to understand what Wittgenstein means when he writes about expression, one needs to reject the assumption that is common to both horns of this dilemma.⁶

III

In the preceding section, I indicated that according to Wittgenstein, a very broad range of self-ascriptions should be understood as expressing the very states or events that they self-ascribe. It is important to add that he does not think *every* psychological self-ascription is such an expression. I might describe myself as angry on the basis of observing (or having it pointed out to me) that my recent behaviour towards my mother seems to be passive-aggressive. Wittgenstein would not hold that such a self-ascription was an expression of anger. He writes: “When someone says ‘I hope he'll come’ – is this a *report* about his state of mind, or an *expression* of his hope? – I can, for example, say it to myself. And surely I am not giving myself a report. It may be a sigh; but it need not.” (PI §58.5).⁷ It is the last sentence of this passage that I want to call to your attention. Whether or not some utterance of “I hope he'll come” should be understood as “a sigh” – that is, an expression – will, according to Wittgenstein, vary from occasion to occasion. This point is developed in the section of the *Investigations* that immediately follows the one just quoted:

The exclamation “I'm longing to see him!” may be an act of expecting. But I can utter the same words as the result of

self-observation, and then they might mean: "So, after all that has happened, I am still longing to see him". The point is: what led up to these words?
(PI §586)

The same sentence that expresses my expectation on one occasion might, on another occasion, report a fact that I have learned via self-observation. If I know my own mental state via observation or inference, then when I self-ascribe it, I do not thereby express it. Now in such a case – one in which I characterize my own state of mind on the basis of self-observation or inference – I do not speak with what philosophers call "first-person authority".⁸ There is, then, a connection between avowing one's state of mind expressively, as it were, and speaking about it authoritatively. My aim in what remains of the present section is to shed light on this connection and thereby bring Wittgenstein's suggestion that we understand avowals as expressing their subject matter into sharper focus.

I shall begin by noting two features of first-person authority. First, if you want to know what I am thinking, planning, hoping, believing, feeling and so on, I am, as a rule, the best person to ask. This is not to say that I am never wrong about my own state of mind, nor is it to deny that there are occasions when another person knows my psychological condition better than I know it myself. I might judge that I am not angry about something when, later, I come to the conclusion that, after all, I was. And a close friend might realize that I am angry before I do. Still, such cases are exceptional. Almost always, the first and best person to ask about someone's state of mind is the subject himself.

A second distinctive feature of first-person authority is that I very often seem able to speak about my own state of mind without basing what I say on any behavioural (or other) evidence. Contrast this with what it takes to speak responsibly about another person's psychological condition. Imagine I tell you that a mutual friend of ours is feeling sad, or that she loves a particular movie, or that she wants to visit Spain. You might simply accept my statement, or you might ask me how I know, or why I believe, what I have said about our friend. Whether or not you ask for grounds, if my claim is responsible, I ought to be able to provide some. I ought to be able to say, for example, "She told me that she's been saving up to buy a plane ticket to Barcelona" or "Well, she looks sad". Now consider a situation in which I tell you that I am feeling sad, or that I love a particular movie, or that I want to visit Spain. Here it would be odd, on the face of it, for you to ask me for grounds or evidence in support of my statement unless you had some quite specific reason for thinking that I might be wrong about myself.

Moreover, it does not seem as if I need to be aware of any behavioural (or other) facts that constitute evidential support in order for me to self-ascribe my mental state responsibly. I can just say, for example, "I am feeling sad", even though I have not exhibited any behaviour that would justify an attribution of sadness to me.

How does Wittgenstein's suggestion that we understand avowals as expressions enable us to better make sense of these two features of first-person authority? First, why am I the best person to ask about *my* state of mind? We might as well ask why mine is the best *face* to consider if you want to know my state of mind. If I am feeling, say, joy or anger, this is liable to be apparent, visible, in my facial expression. And, as we saw in the last section, Wittgenstein holds that the same is true of mental state self-ascriptions: they often make a person's psychological condition manifest. One reason that I am the best person to ask about my state of mind is that asking me is liable to put you in a position to hear in what I say about myself (and so to learn directly and at first hand) what I am thinking or feeling.

Let us turn to the second feature of first-person authority mentioned a moment ago. Why do I not require behavioural (or other) evidence in order to speak about my own state of mind – for example, my own amusement or pain? Well, in order for me to express a pain in my tooth *by wincing*, I do not need epistemic grounds that support the proposition that my tooth hurts. Wittgenstein's point could be put this way: If an expression of pain or amusement takes the form of a sentence about myself, instead of a wince or a laugh, it does not follow that epistemic justification is suddenly called for. If we understand an avowal as akin to a pained wince or an amused laugh, then we should not expect to find that it requires, or is based on, evidential or observational grounds.

IV

Wittgenstein's remarks about the expressive dimension of mental-state avowals have tended not to be taken very seriously by his readers. Instead we find some commentators placing great emphasis on the supposed fact that Wittgenstein's position is here open to easy refutation, while other more sympathetic readers try to minimize the significance of these unfortunate remarks. Both responses are, I suggest, due to a tradition of commentary that reads this strand in his work uncharitably. Or so I shall claim in what follows.

At the end of the preceding section, I emphasized one respect in which (what we might call) the grammar of wincing is different from

that of observation reports: the former do not require epistemic justification; they are not epistemically "based on" observation, inference or anything else. Now, another difference between the grammar of winces and that of reports is that winces do not literally *say* anything either true or false. (I may mislead you by wincing and cradling my cheek when I feel no pain, but if that is all I do, then I have not *lied* to you.) Wittgenstein's oft-repeated suggestion that we understand avowals as, or as akin to, expressions has been widely understood as showing that he is committed to a position according to which an utterance of "I'm in pain" or "I expect an explosion" likewise says nothing either true or false.⁹ Following recent practice, I shall call this sort of position "simple expressivism".¹⁰ Now, simple expressivism is difficult to take seriously. But, I shall argue, Wittgenstein ought not be read as a simple expressivist.

Let us begin by trying to get in view why – or one reason why – simple expressivism is unappealing *prima facie*. Imagine the following scenario: my friend Tom is moving from one apartment to another a few blocks away. Rather than hire movers, he has asked everyone he knows to help with loading and unloading a rented truck. I have agreed to participate, but on the day of the move, I phone him and say: "I'm sorry. I've wrenched my back, and now it's hurting a lot. I just don't think I can help with the move today." Tom, who knows that I dislike lifting large objects, says: "Oh please. Your back doesn't hurt; you just don't feel like getting off the couch." I reply: "No, Tom. I'm telling you the truth; it hurts like hell." Given the sort of expressivism that Wittgenstein is widely thought to espouse, when I say, "I'm telling you the truth; it hurts like hell", I seem to be either mistaken or conceptually confused. For even if I am in awful pain, when I say, "it hurts", I cannot, on this view, be saying something true. Moreover, given simple expressivism, Tom and I should not be understood as even *disagreeing*: he has made a factual claim to the effect that I am not experiencing back pain, and I have made no claim at all. But it seems undeniable that we *are* disagreeing. Hence the unattractiveness of simple expressivism.

It is worth comparing this expressivism about psychological self-ascriptions with a more familiar sort of expressivism about moral discourse, that is, with the sort of "emotivism" championed by, for example, A. J. Ayer (1946) or C. L. Stevenson (1944). An emotivist holds that an utterance such as "It's wrong to eat animals" expresses an attitude of disapproval or a preference but does not say anything true or false. Now, there are striking similarities between this sort of view and the simple expressivism about psychological self-ascriptions

that Wittgenstein is thought to defend,¹¹ but there is a crucial difference as well. Emotivists are moved by a metaphysical conviction that there are no moral facts for moral discourse to be true to. But those who read Wittgenstein as a simple expressivist about psychological self-ascriptions do not take him to be sceptical about psychological facts as such. Rather, Wittgenstein is understood to hold that only *first-personal* uses of psychological predicates yield sentences that are not truth-apt. According to the Wittgenstein that emerges on this reading, if you say that someone other than yourself is unhappy (or in pain or hoping that it will rain ...), you *do* produce a truth-evaluable assertion. This, I suggest, is one reason why the sort of expressivism that Wittgenstein is thought to espouse has proved less attractive than emotivism. Given a familiar set of metaphysical scruples, there is some plausibility in a thought that might be put as follows: "When you and I disagree about whether it is wrong to eat animals, we are *really* disagreeing in our expressed attitudes of approval and disapproval, rather than about any supposed moral facts. I am expressing disapproval of certain practices, while you are expressing approval of them." But no such diagnosis of ordinary disagreement is available to the simple expressivist about psychological avowals. According to simple expressivism, if a friend of mine describes me as "wanting to visit Colorado in the spring" and I correct her, saying, "No; it's Wyoming that I want to visit", she has made a factual claim – a statement that is either true or false – and I have not. There just does not seem to be room here to explain our apparent disagreement by appeal to anything like conflicting expressions of approval and disapproval.

Of course, even if simple expressivism is blatantly unsatisfactory, Wittgenstein might still have defended it. But what speaks against attributing the position to Wittgenstein is not merely that it is unsatisfactory. What speaks against attributing it to him is, first, the absence of any clear or compelling textual grounds for doing so and, second, the fact that reading Wittgenstein as a simple expressivist makes what he *does* say about psychological self-ascriptions substantially less interesting and persuasive. I am not claiming that Wittgenstein ever comes out and asserts that avowals *are* truth-apt. But it is one thing for a philosopher not to explicitly address a question and quite another thing to read him as if he had provided a bad answer to it.

As it happens, Wittgenstein does seem to be committed to allowing that mental-state self-ascriptions have truth-values, even if he does not address the point directly. Consider §136 of *Philosophical Investigations*. Referring to something that he had written in the *Tractatus* ("The general form of propositions is: This is how things are"; *TP* 4.5, Ogden's translation modified), Wittgenstein writes:

At bottom, giving "This is how things are" as the general form of propositions is the same as giving the definition: a proposition is whatever can be true or false. For instead of "This is how things are" I could have said "This is true". (Or again "This is false".) But we have

'p' is true = p
'p' is false = not-p.

And to say that a proposition is whatever can be true or false amounts to saying: we call something a proposition when *in our language* we apply the calculus of truth functions to it.

Wittgenstein goes on to criticize his former self, but he does not reject what is said about "true" and "false" here. He seems to retain a thought that might be put as follows: *given* a proposition, *p*, to which we apply the calculus of truth-functions in our language-game, there is no room to deny that *p* is truth-apt. How does this bear on the question of whether Wittgenstein should be read as a simple expressivist about psychological self-ascriptions? Well, we *do* apply the calculus of truth functions to our psychological self-ascriptions. In reply to someone who has just claimed that no one in the room is experiencing any pain, I might say, "I'm in the room, and I'm in pain, so what you just said is false." Thus, given what he thinks about truth, Wittgenstein seems committed to allowing that mental-state avowals are truth-apt.¹²

Why has Wittgenstein so often been read as holding that psychological self-ascriptions are neither true nor false? Some commentators have failed to see any other way to make sense of the suggestion that we understand avowals as expressions. These readers have clung to an assumption that has no real foothold in Wittgenstein's writings – an assumption that might be stated as follows: "I can let another person know the state of mind that I'm in *either* by expressing it *or* by saying something true about it. But I cannot, in a single speech act, *both* express my state of mind and say, truly, that I'm in it."¹³ If this assumption were correct, then simple expressivism would be entailed by the claim that mental-state avowals are expressions of that which they avow. But the assumption is false.¹⁴ As William Alston (1967: 16) puts the point:

I can express my enthusiasm for your plan just as well by saying 'I'm very enthusiastic about your plan', as I can by saying 'What

a tremendous plan!', 'Wonderful', or 'Great!' I can express disgust at X just as well by saying 'I'm disgusted', as by saying 'How revolting', or 'Ugh'. I can express approval as well by saying 'I completely approve of what you are doing' as I can by saying 'Swell' or 'Good show'. And can express annoyance as well by saying 'That annoys me no end' as by saying 'Damn'.

This shows that expressing and asserting are not mutually exclusive in the way commonly supposed.¹⁵

Once we give up the assumption that expressing and asserting are "mutually exclusive in the way commonly supposed", readings of Wittgenstein as a simple expressivist lose much of their plausibility.

Some of Wittgenstein's readers have taken the fact that he does not like to characterize avowals as "reports" [*Berichte*] (see e.g. *PI* §585) or as "descriptions" [*Beschreibungen*]¹⁶ to suggest that he understands them as neither true nor false. But I think it makes a good deal more sense to hear his reluctance to use these words in connection with psychological avowals as indicating that he takes "report" and "description" to be closely tied to concepts like *observation* and *epistemic justification*, which, as we have seen, he thinks are out of place in this context. Here, consider the following remark from *Zettel*:

To call the expression of a sensation a *statement* [*Behauptung*] is misleading because 'testing', 'justification', 'confirmation', 'reinforcement' of the statement are connected with the word "statement" in the language-game. (Z §549)

According to the kind of reading that I am urging against, this remark might be glossed along the following lines: "A *statement* is the sort of thing that may be true or false. In this passage, Wittgenstein is claiming that when I express a sensation by avowing it, I am not making a statement and so not saying anything either true or false." But two things should, I think, strike us about the quoted passage. First, Wittgenstein does not say that it is *wrong* to call the expression of a sensation a statement (he does not say that the expression of a sensation is *not* a statement); he says rather that it is *misleading* to call such an expression a statement. And, second, the reason why this is misleading has to do with how the word "statement" is connected in our language-game with (i) "testing", (ii) "justification", (iii) "confirmation", and (iv) "reinforcement". Notice what is conspicuously absent from this list: "truth". According to Wittgenstein, it is misleading to refer to an avowal of pain as a "statement" because we think of statements as requiring

justification or confirmation – *not* because we think of statements as having truth-values.¹⁷

V

At the end of §I, I quoted the following from *Philosophical Investigations*: “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” (§246). I noted that this remark might appear puzzling. If I am able to say when I am in pain and when I am not, does it not follow that somehow or other I learn when I am in pain? While we *could* decide to speak of “learning” here, for Wittgenstein this would be to invite confusion. We do not say that I need to learn that I am in pain before I can express my pain by wincing or crying out. Talk of *learning* and of *grounds* or *justification* come together in our language-game. And a way to put some of what has emerged here would be to say that, for Wittgenstein, psychological self-ascriptions are very often outside the logical space of epistemic justification. Sometimes I do need to observe, or to think about, my own behaviour in order to discover, for example, whether I am angry. But, according to Wittgenstein, I am often able to express my state of mind by avowing it, without needing to discover – to learn – anything.¹⁸

Notes

1. John Locke may be read as defending a view of this sort in Book II of Locke (1690) 1975).
2. Editions referred to in this chapter are *Philosophical Investigations* (2001a) and *Zettel* (1967).
3. I have departed from the published English versions of both of these passages by translating Wittgenstein’s “*Äußerung*” as “expression”. (The published English translations of Wittgenstein’s late writings vacillate between translating “*Äußerung*” as “expression” and as “manifestation”.)
4. See note 17.
5. The phrase “behaviourist in disguise” is from *Philosophical Investigations* §307. See also the end of §244 (which I quoted at the start of §II), where Wittgenstein’s interlocutor hears him as suggesting that “the word ‘pain’ really means crying”.
6. A philosopher who is in the grip of the interlocutor’s dilemma has limited resources for understanding psychological self-ascriptions. Wittgenstein represents the situation of such a philosopher in the following passage from *Zettel*: “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. – 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 of the real expectation, and was that a special mental process?” (§53).

Wittgenstein aims to show that if it appears to us that our only options for understanding avowals of (say) expectation are: (i) as remarks about behaviour, or (ii) as descriptions of hidden mental processes of which observable behaviour is merely a “side-effect”, then (to quote the section in the *Investigations* that immediately follows the one in which Wittgenstein’s interlocutor accuses him of being “a behaviourist in disguise”) “the decisive movement in the conjuring trick has [already] been made” (§308).

7. I have again translated “*Äußerung*” as “expression”. (See note 3.)
8. I am simplifying things a bit. It sometimes (often, I think) happens that a mental-state self-ascription is to some extent an expression of that which it self-ascribes and to some extent a report based on evidence. In such cases, we may be said to speak with a degree of first-person authority. (For more on this point, see Finkelstein 2003: 122–6.)
9. Wittgenstein is read this way by, e.g., Tomberlin (1968: 91); Hacker (1986: 298); Fogelin (1987: 197); Rosenthal (1993: 203); and Bar-On & Long (2001: 321–2).
10. Dorit Bar-On (2004) distinguishes between “Simple Expressivism”, according to which avowals are understood to be expressions of their subject matter and so not truth-apt, and “Neo-Expressivism”, according to which an avowal may both express its subject matter and be true.
11. And a defender of simple expressivism about psychological self-ascriptions would encounter difficulties that closely correspond to familiar problems faced by expressivists about moral discourse.
12. This paragraph is indebted to Jacobsen (1996).
13. Thus David Rosenthal writes: “I can communicate my suspicion that the door is open either by expressing my suspicion or by explicitly telling you about it ... In saying I suspect something, I report, rather than express, my suspicion” (Rosenthal 1993: 200).
14. Consider the remark by Rosenthal that I quoted in the last note. Once someone has sincerely asserted, “I suspect that my wife is having an affair”, could he then be rightly described as “never having expressed a suspicion that his wife was having an affair”? No; one common way to express a suspicion is by *saying* that one is suspicious.
15. Alston himself holds that non-linguistic expressions such as “sneals, looks, and tones of voice do not express feelings in anything like the sense in which they are expressed by linguistic expressions [regardless of whether the latter happen to be assertions]” (Alston 1967: 17). His position thus turns out to be very different from Wittgenstein’s.
16. Although see *Philosophical Investigations* §290, which shows that Wittgenstein has no real objection to our characterizing an avowal of pain as a “description”, as long as we manage to bear in mind how different the language-game of describing one’s own mind is from that of describing, say, one’s room.
17. In §II, I quoted a passage from *Zettel* that includes these lines: “Sentences in the third person of the present: information. In the first person present: expression. (Not quite right.) // The first person of the present akin to an expression” (Z

§472). Given that Wittgenstein very often characterizes psychological avowals as expressions, why does he qualify the point here? While I cannot be entirely certain, I can suggest an answer. The two sentences that precede "Not quite right" might give the impression that Wittgenstein thinks mental-state avowals are more like cries, winces and moans than he takes them to be. It is fairly natural to read that pair of sentences as expressing a thought that could be put as follows: "If I tell you, for example, that my sister has a headache, I am offering you a bit of information, a fact. But if I say 'I have a headache', I am not stating a fact – nor saying anything true – but merely, as it were, wincing with words." I believe that it is this sort of misreading of him that Wittgenstein has in mind when he worries about the pair of sentences that precede "Not quite right" in the quoted passage.

18. Thanks to Thomas Lockhart and Joshua Scodel for helpful comments on a draft of this chapter. For more about this way of reading Wittgenstein on avowals and expression, see Finkelstein (1994, 2001, 2003).

Chronology of Wittgenstein's life

- 1889 Born on 26 April in Vienna, Austria.
 1908 Began studying engineering at Manchester; reads Russell's *Principles of Mathematics*.
 1911 Visits Frege (Frege refers him to Russell); moves to Cambridge and meets Russell, attends Trinity College at Cambridge and studies with Russell.
 1912 Review of P. Coffey, *The Science of Logic for Cambridge Review*.
 1913 Dictates *Notes on Logic*; moves to Norway.
 1914 Dictates notes to G. E. Moore in Norway.
 1914–18 Serves in Austrian army in the First World War and is captured in October 1917; held near Monte Cassino, Italy; completes *Tractatus Logico-Philosophicus* in August 1918.
 1919–28 Leaves philosophy; studies at a Vienna college for elementary school teachers; teaches elementary school; works as a gardener at a monastery; designs a house in Vienna for one of his sisters; publishes *Tactatus* in 1922.
 1929 Returns to philosophy at Cambridge; publishes "Some Remarks on Logical Form", *Proceedings of the Aristotelian Society*.
 1930–35 Becomes a Fellow of Trinity College. Dictates *The Blue Book* and *The Brown Book*.
 1936 Lives for over a year in Norway and begins *Philosophical Investigations*.
 1938–9 Returns to Cambridge and succeeds Moore as the chair in philosophy.
 1940–44 Serves as a porter at Guy's Hospital in London; works in a medical lab.
 1945–7 Lectures and teaches at Cambridge.
 1949 Visits Norman Malcolm in America (at Cornell).
 1949–50 Returns to England; lives in Oxford.
 1951 Dies on 29 April in Cambridge.
 1953 *Philosophical Investigations* is published.