CHAPTER 1

Anatomy of a Muddle: Wittgenstein and Philosophy

Alexander George

> It is not sufficiently considered, that men more frequently require to be reminded than informed.
>
> — Samuel Johnson

0 Ludwig Wittgenstein has a problem with philosophy. This claim will hardly generate dissent. Though perhaps it should: for it might be objected that Wittgenstein is leery of generalizations, including ones about what troubles philosophy. He emphasizes that there is not just one way of going astray in philosophy (PI §133).

Nevertheless, Wittgenstein does have a recognizable approach that he regularly pursues in his philosophical investigations. There is a problem that he often presses, a form of criticism that he often develops, against traditional pursuits of philosophy, though in any given case its true force only becomes clear in the particularities of its execution.[1] It is surprisingly difficult to say clearly what this problem is. But it is worthwhile to try, for not only is this criticism a hallmark of his thought but it is closely connected to other central features of it, for instance, to his conception of language and of the nature of philosophical investigation. These features can be properly understood only in concert with a correct view of his terms of criticism of traditional philosophy. In what follows, I shall articulate a problem Wittgenstein sees with philosophy, show how it illuminates otherwise peculiar features of his investigations, and finally consider an illuminating way in which his goals might be thwarted.

1.1 It will be helpful to begin by briefly deploying the *calculatus eliminatus* to explore what Wittgenstein's criticism is *not.* Most obviously, Wittgenstein

[1] The general terms of criticism that I describe, and the correlative approaches Wittgenstein develops to press that criticism, are important *leitmotifs* of his entire later thought; whether this is the only form of problem he presses does not concern me here.

does not aim to argue for the falsity of a philosophical proposition. If one knows anything about Wittgenstein's thought, one knows this. Of course Wittgenstein is not the only person in the history of philosophy who eschewed judgments of falsity when rejecting a philosophical pronouncement: David Hume, to take but one example, cannot always be read as arguing for the falsity of some opponent's claim. But Wittgenstein, throughout his life, was adamant that the proper critical stance to take toward traditional philosophical claims is *not* that they are false.

Even those with a passing knowledge of Wittgenstein's philosophy will know that his terms of criticism involve the notion of nonsense. But what is nonsense? It is natural to think that Wittgenstein's problem with the philosopher's claim is that it cannot be made sense of because it is couched in a way that somehow violates rules of his language. One might elaborate this thought as follows: claims are made within language, languages employ words that have meanings, and those meanings may occasionally conflict in such a way as to render the entire claim itself unintelligible, that is, in such a way as to render it nonsense. This view may go hand in hand with a view of Wittgenstein as a language policeman, someone who tasks himself with discovering the meanings that guide our use of language and stopping those who do violence to those meanings through illegal combinations of expressions.

This view is not tenable, however, in light of Wittgenstein's regular insistence that he has no objection to the form of words per se used by the philosopher: "It is not our intention to criticize this form of expression," he writes (BB 7). Indeed, Wittgenstein might have no problem with those very same words as spoken by someone else. Whatever Wittgenstein's criticism is, it is not directed just at a "form of expression."

It might be thought that keeping in mind Wittgenstein's focus on the *use* of expressions (as opposed to meanings allegedly attached to them) will be enough to get us back on track. But even with this in view, there are ways of misunderstanding Wittgenstein's critique. For instance, if we recast Wittgenstein's policing as targeting circumstances of use and say that his complaint with the philosopher is just that she uses words in unordinary ways, then we will again not have matters quite right. For Wittgenstein has no problem with familiar locutions being used unfamiliarly. Indeed, a distinctive feature of his method is precisely to conjure up circumstances in which familiar expressions are used differently: his intermediate cases and imagined peoples all involve scenarios in which common expressions are used in alternative ways.

To sum up, we have been considering attribution of the following line of thought to Wittgenstein:

(1) To employ an expression in an unordinary way is to utter nonsense.

(2) Philosophers often employ expressions in unordinary ways.

(3) Hence, philosophers often utter nonsense.

We have just been noting, however, that (1) is at best questionably attributed to Wittgenstein. From the fact that someone employs an expression in an unusual way, Wittgenstein would not conclude that she has lapsed into nonsense. If someone says "I feel the visual image to be two inches behind the bridge of my nose," a form of expression that is alien to us, Wittgenstein will not judge that he "is telling a lie or talking nonsense"; rather, he simply says, "this phrase has sense if we give it sense" (BB 10, 7). Wittgenstein's problem with philosophical claims is not most illuminatingly put by saying that they employ expressions in unusual ways.[2]

1.2 I shall turn now to a positive characterization of Wittgenstein's view. It will be helpful to consider an example of a situation in which Wittgenstein's antennae might begin to quiver. Imagine that someone expects that Bill will attend the party. Now imagine further that she says that she has a certain feeling in her stomach. Should we describe her feeling as that of "expecting that Bill will attend the party"? (Wittgenstein attributes such a view to Bertrand Russell.) It might come as a surprise that Wittgenstein would have no objection to so describing that feeling. Nevertheless, he would immediately caution, it should not be thought that by doing so we have a clear understanding of what "the feeling of expecting that *x* will attend the party" means for any given value of "*x*." It is true that a particular feeling has been baptized with a complex name, but the complexity of the name has been given no role in determining to which feeling we are referring. So while there is nothing wrong in itself with dubbing a sensation as "the feeling of expecting that Bill will attend the party," there may be a temptation down the road to assimilate this expression to, say, "the likelihood that Bill will attend the party." And if we do so, then we will be inclined to treat the parameters of the latter expression (for instance, "Bill") as present also in the former.[3] That is, because we can talk about "the likelihood that Bettina will attend the party," we will be

[2] There is something else problematic about attributing premise (1)—"to employ an expression in an unordinary way is to utter nonsense"—to Wittgenstein, if it encourages the thought that he takes there to be a determinate totality of uses that constitutes "the ordinary use" of an expression. I return to this issue in Section 1.3 and, again, in Section 2.3.

[3] In the *Blue Book*, Wittgenstein calls these parameters—for instance, "Bill" in the expression "the likelihood that Bill will attend the party"—"arguments" (BB 21).

inclined also to speak of "the feeling of expecting that Bettina will attend the party"; and this is an expression about whose use we are none too clear.

Before exploring the kind of confusion Wittgenstein thinks such assimilations can lead to, it is worth noting that warnings about such a slide—whereby an innocent-looking notation induces a false sense of the intelligibility of some parametric variation—reappear in central texts of analytic philosophy by thinkers who were no doubt influenced by Wittgenstein.

In "Two Dogmas of Empiricism" (1951), Quine considers an objection to his doubts about analyticity. He imagines someone who protests that analyticity is messy to characterize merely because ordinary language is messy. By contrast, we are assured, if we confine our attention to sharply defined artificial languages then the notion of *analytic truth* can be explained as clearly as one pleases. Quine's difficulties in characterizing the analytic truths of ordinary language, according to this objection, stem not from the notion of analyticity but rather from our lack of understanding of the structure of ordinary languages. Following Carnap, for instance, we can articulate a set of "semantical rules" of a given formal language that will settle which of its formulae count as analytic truths. For instance, assume that we are considering a particular formal language, *L*, and that we have specified what its semantical rules are. Then, by appeal to these, we can specify a certain subset of its formulae, the set of all and only those formulae that are analytic in *L*. What, Quine's interlocutor continues, could possibly be objectionable about this?

I think Quine would agree that so far there is nothing objectionable at all. A certain subset of the formulae of *L* has been clearly defined and if one wishes to call the formulae in that subset the analytic truths of *L*, one should feel free to do so. The problem comes when one thinks that the relation __ *is analytic in* __ has been thereby illuminated in any way. Such illumination would require that we treat "*L*" in "analytic in *L*" as a parameter, but nothing has been said so far to give us a handle on, say, the property "analytic in *M*." Quine puts the point this way:

> We may, indeed, view the so-called rule as a conventional definition of a new simple symbol "analytic-for-L_o," which might better be written untendentiously as "*K*" so as not to seem to throw light on the interesting word "analytic". . . . By saying what statements are analytic for L_o we explain "analytic-for-L_o" but not "analytic," not "analytic for." We do not begin to explain the idiom "*S* is analytic for *L*" with variable "*S*" and "*L*," even if we are content to limit the range of "*L*" to the realm of artificial languages. (Quine 1951, 33–4)

Again, an innocent notational predilection may lead us astray in giving us the impression that we have endowed a range of other locutions with significance.

In 1956, Wilfrid Sellars first presented his paper "The Myth of the Given: Three Lectures on Empiricism and the Philosophy of Mind." One of his targets was contemporary and classical sense-datum theories. At one point he considers the possibility that

> someone might introduce so-called sense-datum sentences as code symbols or "flags," and introduce the vocables and printables they contain to serve the role of reminding us of certain features of the sentences in ordinary perceptual discourse which the flags as wholes represent. In particular, the role of the vocable or printable "sense datum" would be that of indicating the symbolized sentence contains the context ". . . looks . . ., " the vocable or printable "red" that the correlated sentence contains the context ". . . looks red . . . " and so on. (Sellars 1956, 27)

Now Sellars has no objection to so introducing sense-datum talk. Indeed, doing so might be useful for various purposes. The problem is that it is easy to forget how this code was introduced and as a consequence easy to mistake the code elements for the identical strings of symbols in everyday talk. The philosopher, focused on the form of the expression instead of on how it was introduced and is used, easily confuses the two. And from there, it is easy to convince oneself that these code statements have an independent life and indeed involve concepts that help to clarify or explain our everyday discourse about how the world appears to us. Sellars thinks that this is one route down which philosophers have traveled to the idea of an intelligible and explanatory sense-datum language. As Sellars puts it:

> It would take an almost superhuman effort to keep from taking the vocables and printables which occur in the code . . . to be *words* which, if homonyms of words in ordinary usage, have their ordinary sense, and which, if invented, have a meaning specified by their relation to the others. (Sellars 1956, 29)

Absent this superhuman effort, "one may be tempted to try to eat his cake and have it" by both treating "sense-datum talk as *merely another language*" and taking it to "have an explanatory function" (1956, 30).

We can trace back this idea even further if we recall a book that Wittgenstein, Quine, and Sellars all read: the *Tractatus.* For, as Wittgenstein himself notes in the *Blue Book* (BB 21), a version of this idea makes an appearance there. In the context of a discussion of the Tractarian claim that

all propositions are truth-functions of elementary propositions, Wittgenstein writes:

> The arguments of functions are readily confused with the affixes of names. For both arguments and affixes enable me to recognize the meaning of the signs containing them.
>
> For example, when Russell writes "$+_c$," the "$_c$" is an affix which indicates that the sign as a whole is the addition-sign for cardinal numbers. But the use of this sign is the result of arbitrary convention and it would be quite possible to choose a simple sign instead of "$+_c$"; in "$\sim p$," however, "p" is not an affix but an argument: the sense of "$\sim p$" *cannot* be understood unless the sense of "p" has been understood already. (TLP 5.02)

Though the details are unimportant here, Wittgenstein goes on to formulate an objection to Frege in terms of this distinction, namely that his "theory about the meaning of propositions and functions is based on the confusion between an argument and an affix."[4] Using this terminology, we can say that in "the feeling of expecting that Bill will attend the party," the name "Bill" functions not as an argument but as an affix.

But in fact, this kind of concern can be traced back yet further and, again, to a source that was certainly read by Wittgenstein, Quine, and Sellars: Frege himself. After arguing that, for instance, *being two in number* is not a property of an object but rather of a concept (in his technical sense of "object" and "concept"), Frege in *The Foundations of Arithmetic* (1884, §55) proceeds to offer a recursive definition of the property *being* ***n*** *in number*. To say that F is 0 in number is simply to say that there exists nothing which is F:

$$(\exists_0 x)Fx =_{\text{df}} \neg(\exists x)Fx$$

To say that F is exactly 1 in number (more colloquially, that there is exactly one thing that is F) is to say that at least one object is F and furthermore that the number of things that are F and not identical to that object is 0:

$$(\exists_1 x)Fx =_{\text{df}} (\exists x)(Fx \wedge (\exists_0 y)(Fy \wedge x{\neq}y))$$

And likewise, to say that there are exactly 2 Fs is to say that at least one thing is F and furthermore there is exactly one object that is F and not identical to that thing. It can thus be defined as:

[4] For a brief discussion of Wittgenstein's objection to Frege's theory of propositions and functions, see Black 1964, 238–40.

$$(\exists_2\, x)Fx =_{\text{df}} (\exists x)(Fx \wedge (\exists_1\, y)(Fy \wedge x{\neq}y))$$

More generally, we can define what it is for exactly $n{+}1$ things to be F:

$$(\exists_{n+1}\, x)Fx =_{\text{df}} (\exists x)(Fx \wedge (\exists_n\, y)(Fy \wedge x{\neq}y))$$

In this way, using just the apparatus of first-order quantification theory and the identity relation, Frege was able to define all the numerically definite quantifiers (see Frege 1884, §55).

Frege was not content with this analysis. But it is important to note that his dissent had nothing to do with the account's correctness: the logical structure of, say, "There are exactly 2 Fs" is exactly as specified by this recursive definition. The problem for Frege was rather that this notation misleads in giving the impression that the definition tells us something about the number 2 or about any particular natural number (which Frege announced as his goal at the outset of *Foundations*). To put the point formally, the notation suggests that the numeral "2" contributes semantically to the meaning of the claim that "There are exactly 2 Fs" just as the name "Hume" contributes semantically to the meaning of "Hume is Scottish." If we were to unpack the truth conditions of the latter statement, we would perforce make reference to Hume. However, if we unwind the truth conditions of "$(\exists_2\, x)$Fx," we will not find ourselves referring to the number 2. For that reason, Frege thought that the numeral "2" in "$(\exists_2\, x)$Fx" does not function as it does in, say, the statement "2 is an even prime number." To use the *Tractatus'* terminology, the numeral "2" in "$(\exists_2\, x)$Fx" does not function as an argument. Given that the *Tractatus* introduces this notion in order to criticize Frege, there is some irony in the fact that Frege himself anticipates that very concern about notation. "It is only an illusion," Frege writes, that "2" in "There are exactly 2 Fs," as defined earlier in this section, functions as it does in "2 is an even prime number," an illusion fostered by our terminology.[5]

Let us now return to Wittgenstein and to a comical example of how our notation can induce such "illusions." At one point, he asks us to imagine someone who, for whatever reason, wishes to call a non-painful form of tooth decay "unconscious toothache." Again, one might think that

[5] Frege 1884, §56: "It is only an illusion that we have defined 0 and 1; in truth we have only determined the sense of the phrases

'the number 0 belongs to,'
'the number 1 belongs to';

but this does not allow us to distinguish 0 and 1 here as independent, reidentifiable objects."

in itself this is troubling to Wittgenstein as these words are now being used in an unusual way. But this is not so: "There is nothing wrong about it," he insists (BB 23), "as it is just a new terminology and can at any time be retranslated into ordinary language." We might say that the term "toothache" does not function here as a parameter; it is not an argument but an affix, in the *Tractatus*' terminology. "Unconscious toothache" is not an instance of "unconscious *x*." That is to say, the complex expression has not been explained to us in such a way that we know how to use it in ways analogous to the ways in which we might use the superficially similar "unconscious pregnancy," "unconscious cancer," etc.

How the expression was introduced can easily be forgotten, however, and we may find ourselves transforming "Smith has an unconscious toothache" in ways that we readily transform "Smith has an unconscious pregnancy." For instance, just as we might move from the latter to "Smith is pregnant but does not know it," so we might be tempted to move from the former to "Smith is in pain but does not know it." This, however, is a sentence for which we have no handling.

This kind of slide can be well illustrated by considering how a comparable situation might arise from Frege's account of the numerical quantifiers. Since we readily infer from "2 is an even prime number" that "There is at least one even prime number," the "illusion" of parity with "There are exactly 2 Fs" might encourage us to infer "There is at least one thing such that there are exactly that many objects that are F" or "$(\exists y)(\exists_y x)Fx$." We might, that is, treat "2" as functioning in the same way in both sentences, as occupying an argument position that can be existentially generalized. Now as it turns out, we can assign a meaning to this last claim. But that understanding has to be given to us independently: there is nothing in the recursive definition of the numerical quantifiers that teaches us how to handle this claim. As Frege puts it (1884, §66), our understanding of such new expressions is "no thanks to" his definition. (This independent understanding is precisely what Frege aims to provide in the remainder of his *Foundations*.)

1.3 We are now in a better position to articulate what the difficulty is that surrounds "Smith has an unconscious toothache," according to Wittgenstein. As we have seen, it is not the expression itself or the novel use assigned to it. The full problem is rather our inability to keep this use in focus. For we have a tendency to forget our initial resolve as to its use and instead to imagine that its use is of a piece with that of "Smith is pregnant but not conscious of that fact." To deploy the *Tractatus'* terminology, we

had resolved to treat "toothache" as an affix, but we easily drift unawares into treating it as an argument.

As we lose sight of the use on which we had previously settled, we come to see the expression through the lens of syntactically analogous but in fact unrelated locutions. Wittgenstein sketches a number of different ways in which we might then react. We might dub the first that of the bold explorer: he will treat the existence of "unconscious toothache" as a major discovery about the mind and "he will say it like a man who is destroying a common prejudice" (BB 23). Or one might react as a skeptic and "deny the possibility of unconscious toothache" (BB 23). Or one might respond as a theoretician, "perhaps ask such a question as 'How is unconscious toothache possible?'" (BB 23), and then seek to develop a theory to accommodate this strange phenomenon. Or one might find oneself oscillating among all these responses, and yet others.

What these reactions have in common is that they presuppose that the words in question express a claim along the lines of someone's being pregnant without knowing it. But the fact is that we have not yet settled how to use the expression "unconscious toothache" along the lines of "unconscious pregnancy," and the only use we have given it so far—namely that of "painless tooth decay"—is one of which we have now lost sight. The speaker is left with a form of expression that she is convinced is significant but whose significance has yet to be resolved. "The phenomenon," Wittgenstein says, "is that of *irresolution*" (VW 235). Here, the speaker neither has in mind a definite use that has been marked out ("painless tooth decay") nor appreciates that formal analogies ("unconscious pregnancy") will not suffice to settle some use along other lines. This is an unwitting irresolution: the speaker thinks her words have a role in the language when in fact that role is still quite up in the air. Later, Wittgenstein will put this point by saying that the "confusions which occupy us arise when language is like an engine idling, not when it is doing work" (PI §132); idling, he might have added, while we believe the gears to be fully engaged.[6]

Before turning to a different kind of example of such irresolution, it will be worthwhile to make a number of quick observations. First, Wittgenstein speculates throughout his work on the forces that feed such irresolution. One that is of great interest to him is the assumption that

[6] While writing this essay, I had occasion to reread Stanley Cavell's "The Availability of Wittgenstein's Later Philosophy." It seems to me there is a congruence between my understanding of Wittgenstein and Cavell's suggestion that for Wittgenstein "the philosopher has no position at all" (Cavell 1962, 83).

words have determinate meanings that accompany them from one circumstance of use to another. We naturally assume that an expression contributes to the significance of a sentence in a manner that is systematically related to the way it contributes to the significance of any other sentence in which it might appear. There are several reasons why this assumption about meaning leads to trouble, according to Wittgenstein, but in this connection it is the fact that it encourages the philosopher to think that in order to endow her words with meaning no further work on her part is required beyond concatenating them in familiar ways. Thus, after Wittgenstein expresses perplexity about the sentence "I feel in my hand that the water is three feet under the ground," he imagines the response: "Surely you know what it means. You know what 'three feet under the ground' means, and you know what 'I feel' means!" To this, Wittgenstein replies: "I know what a word means *in certain contexts*" (BB 9). This sentence, he says, "combines well-known words, but combines them in a way we don't yet understand. The grammar of this phrase has yet to be explained to us" (BB 10).

Secondly, Wittgenstein often suggests that we regularly misinterpret our confusion about how to handle the relevant expression as the product of our grappling with philosophically deep issues. The confusion, which is actually a consequence of our irresolution about how to handle our expressions, is projected by us onto what we think the expressions are about and thereby misconstrued as an indication of the depth of the subject matter (relatedly, see PI §111). A muddle about meaning is mistaken for a problem of philosophy.

Thirdly, it is worth reemphasizing how this criticism differs from that which is often attributed to Wittgenstein. As remarked earlier, Wittgenstein is frequently interpreted to hold that the philosopher goes astray when she uses words in a fashion that conflicts with their ordinary usage. For him, it is said, the way the philosopher insists on applying her words clashes with the ordinary way in which they are used and thus ought to be used. But there are a number of problems with this construal. For one thing, it is misleading to say that Wittgenstein believes there is such a thing as "the ordinary way" in which words are used: his explorations reveal highly varied and perhaps even uncircumscribable ways in which we put our language to use. There is no determinate totality of uses that constitutes "ordinary usage." For another, we have seen that Wittgenstein does not believe that in matters of language use *is implies ought*: even if there were a definite canonical fashion in which we deploy our words, there would be no requirement that we stick to it. And finally, there is the point