

A COGNITIVE SELF-THERAPY

PI 138–97

Eugen Fischer

Wittgenstein compared his treatment of philosophical questions to the cure of an illness, his philosophical methods to different therapies. In this paper, I try to spell out the point of these comparisons. To this end, I analyse Wittgenstein's problems and proceeding in sections 138–97 with the help of some new concepts, in part adapted from clinical psychology, namely, Aaron Beck's 'cognitive therapy'. I first use them to conceptualise the problems at issue in such a way as to bring out why anything worth calling a 'therapy' is required, in the first place. I then employ the model of cognitive therapy to clarify what Wittgenstein is doing in response. This will familiarise us with a little noted but highly important kind of philosophical predicament, and with a straightforward approach to it that is in many ways revolutionary.

1 Introduction

In section 138 Wittgenstein turns to a problem about sudden understanding, which he apparently comes to terms with in section 197, the last section of the *Investigations* that was contained already in the early version of the work (the *Frühfassung* of 1937–8, TSS 220–1, henceforward: FF). In between, Wittgenstein repeatedly returns to the theme of the problem, while moving criss-cross through a bewildering range of topics: he discusses first apparently sudden understanding of the word 'cube' (139–42), then understanding – not of another expression but – of the system of a number series (143–50) and, in particular, its sudden understanding (151–5). He digresses into a discussion of reading, conceived of in a rather unusual manner as the transformation of written signs into sounds or *vice versa* (!), with or without understanding (*sic*) (156–78), then returns to his previous topic (179–84), considers next how the way in which a formula is meant determines a number series (185–90) and, finally, returns once more to the topic of sudden understanding, though not of

such a series but of an (unspecified) 'word' (191–7), with a digression into the way in which a machine symbol determines a machine's movements (193–4). The apparent marks of ineptitude noted in the introduction to the present volume are thus very much in evidence: Wittgenstein moves swiftly through various topics that appear to be, at best, rather loosely related, digresses twice into topics of no evident philosophical relevance, and keeps coming back to one main problem, whose discussion he might well seem to have been unable to integrate into one sustained treatment. The present chapter is to reveal the system that underlies this apparent chaos: not a systematic doctrine but a methodical approach of philosophical therapy, in whose light all the apparent marks of ineptitude turn out to make perfectly good sense.

Before we try to clarify what Wittgenstein did here, and why, we might as well point out what he was *not* trying to do: he was not much concerned with advancing general and non-trivial theses, claims of the sort philosophers typically make and often argue for. Consider what is generally regarded as his discussion of the notion of understanding: sections 143–55 and 179–84. These have frequently been taken to advance general claims about criteria of linguistic understanding. If taken at face value, however, practically all of the sections thus interpreted make far more restricted claims: claims, e.g. about the use of the expressions 'Now I know how to go on' and 'Now I can go on' with respect to number-series (180a1).¹ A certain idea recurs in at least two such remarks, namely, the idea that the use of various things is *one* criterion for some sort of understanding: the use one makes of a picture is a criterion of one's understanding of a method of projection (141c); and one's application of an algebraic formula in generating a number-series is one criterion of how one understands that algebraic expression (146b6). But while clearly implying that the same point is recurring here (146b4), Wittgenstein nowhere bothers to explicitly sum up what one might take to be the general idea, in a neat assertive statement he would then go on to endorse. Indeed, the entire discussion contains only three non-trivial remarks of suitable generality: 150, 182 and 154d. The first of these is so extremely vague as to represent rather a pointer than the expression of a statement. The second, only added at the rather late stage of the *Zwischenfassung*, is primarily concerned with setting a 'grammatical exercise' rather than advancing the (at any rate rather modest) 'grammatical claim' that the criteria we accept for 'fitting', 'being able', and 'understanding' are 'more complicated' than one is first inclined to think. This leaves us with one single remark that looks like the expression of a typical philosophical claim (154d) – and even this was only added in the process of revision (in handwriting, in BFF, i.e. TS 239). Advancing philosophical theses, 'material' or 'grammatical' in nature, thus seems to have been, at best, a rather low priority in Wittgenstein's 'discussion of understanding'.

So what was he trying to do, instead? This chapter will explore a pleasingly pedestrian answer: in the investigation of sections 138–97, he was trying

to cope with the problems that he explicitly set out in its course, chiefly in sections 138–9, 151–3 and 191–7. A proper understanding of the nature of these problems will prove the key to an understanding of Wittgenstein's efforts to cope with them. With the help of a few new concepts, I will therefore analyse in some detail how these problems arise, and of what kind they are. This will allow us to make good sense, first, of Wittgenstein's declared aims and, second, of the approach he employed to attain them. In both cases, the crucial insights will be gained by bringing out precisely those aspects of his problems and proceeding that lend a point to his comparison between the treatment of a philosophical question and that of an 'illness' (255), and of his methods with 'different therapies' (133d). Which is all to the good, as these comparisons, integrated only into the final version of the *Investigations*, arguably represent Wittgenstein's final stage of insight into the nature of his own work.²

2 Some new notions

A radio-astronomer is rumoured to have knocked his instruments until they stopped emitting noises he found unintelligible and therefore put down to a malfunction. He was more attentive than others who simply failed to notice what they were not prepared to understand. The pulsars were then discovered by yet others who picked up the same signals when already possessing some concepts to make sense of them. The *Investigations* are replete with features that are, at first sight, positively odd. Some of them occasioned some violence to the text (for a start, at the hands of the translator), others quietly went unnoticed – both, arguably, also for lack of handy concepts to make sense of them. I wish to remedy this lack, though talk of 'introduction of concepts' may be a little grand for the occasion: although some of them were first developed by clinical psychologists and have a technical ring to them, the terms in question are very convenient to bring into view perfectly everyday phenomena ordinarily talked about in a roundabout way. And this is how we shall use these terms: to bring into view some phenomena we are actually quite familiar with but are all too willing to forget about the moment we style ourselves 'rational thinkers' (rather than trying to be, more plainly, reasonable people).

In everyday life, we exercise varying degrees of control over our thinking: sometimes we make a conscious effort, e.g. to apply certain rules of inference, sometimes we simply let our thoughts drift or think of nothing in particular. In the latter case, we may suddenly be hit by a thought that spontaneously occurs to us out of the blue ('I forgot to turn off the stove!'). Occasionally, such a thought recurs even once we have reflectively dismissed it ('Don't be silly! You always check before leaving the house'). Psychologists speak of more or less 'controlled' and 'automatic' thinking, and of 'autonomous thoughts':

A subject's *thought* is *autonomous* iff it spontaneously occurs to the subject when he is exercising a low degree of control over his thinking, and recurs, in the same form or as a variation on the same theme, even once it has been reflectively dismissed.

While some such thoughts are highly salient ('The stove!'), others are not, in particular when they are at odds with our own considered beliefs, which tend to inspire conjectures about what thoughts we 'must be having'. In this case, autonomous thoughts may be sincerely reported only the moment the subject makes a conscious effort to attend to what he is 'tempted' or 'inclined to say', whether or not he judges the idea in question reasonable. This happened, e.g. in the case that first inspired the cognitive approach to emotional disorders (Beck's 'Anna O.');

A woman who felt continuous unexplained anxiety in the therapy sessions was describing certain sensitive . . . conflicts. Despite mild embarrassment, she verbalized these conflicts freely and without censoring. It was not clear to me why she was experiencing anxiety in each session, so I decided to direct her attention to her thoughts about what she had been saying. Upon my inquiry, she realized that she had been ignoring . . . [these thoughts]. She then reported the sequence: 'I am not expressing myself clearly . . . He is bored with me . . . He will probably try to get rid of me.' [Thus] . . . her chronic anxiety during the therapy sessions began to make sense. Her uneasiness had nothing to do with the . . . conflicts she had been describing. But her self-evaluative thoughts and anticipations of my reactions pointed to the essence of her problem. Even though she was actually quite articulate and interesting, she had continual thoughts revolving around the theme of her being inarticulate and boring.

(Beck 1976: 32)

These thoughts are part and parcel of a more comprehensive habit of thought: frequently, they are interpretations of remarks made or events observed (say, the tired therapist's suppressed yawn). These interpretations are usually wrong, and not even natural in the circumstances. But they are reasonable in the light of assumptions that are specific instantiations of the general theme in question (here: 'As people realise, I am inarticulate and boring'). Such misinterpretations usually go with inferences that tacitly presuppose assumptions articulating this very theme, i.e. inferences that are invalid as they stand but can regularly be rendered valid by supplying such an assumption as a further premise. Even though otherwise a perfectly competent thinker, the subject thus regularly makes leaps of thought, of two kinds: misinterpretations and invalid inferences, informed by assumptions

revolving around always the same theme. They occur mostly in automatic but also in more controlled thinking, indeed, even in quite careful reflection. The habit of making such leaps of thought is 'autonomous' in the sense that the subject is not aware of relying on the assumptions in question and may, indeed, even reflectively reject them as false or unwarranted, all along. These tacit assumptions may, but need not, occur themselves to the subject as autonomous thoughts. Thus,

an autonomous cognitive habit is the habit of making leaps of thought informed by assumptions instantiating or articulating a certain theme, assumptions the subject is not aware of relying on and may even reflectively reject.

While it is in (cognitive) psychotherapy that most trouble is taken to identify such habits, they are more or less evident in the thinking of every human being: all of us frequently leap to conclusions that rely on tacit assumptions we are able to identify only with some effort. (Philosophers will be familiar with the fact that it is significantly more difficult to identify tacit premises in others' arguments when one shares them oneself, equally tacitly.) Some such assumptions recur systematically. And presumably some of these assumptions or more specific implications of them would occur to us as autonomous thoughts, if (like Beck's patient) we attended to what we are inclined or tempted to say, when making those leaps of thought.

Finally, Beck's example illustrates nicely that autonomous thoughts may have much the same emotional consequences as considered beliefs: the lady felt (almost) as anxious in the session as if she had sincerely and reflectively believed she was boring her therapist. (Before you find this outrageous, recall that the thought that you left the stove on may spoil your afternoon stroll, even once you have thought the matter over and concluded that you did not.) This is of the essence of emotional disorders. They involve feelings that are, in a quite specific sense,

pathological: they are intelligible in the light of autonomous thoughts but utterly unreasonable in the light of the beliefs the subject reflectively holds throughout.

This term is, of course, more frequently used in other senses. But this one is arguably the most pungent. After all, there need not be anything wrong with you if you have feelings that are 'particularly severe and disabling' or 'utterly unwarranted by the objective situation': severe and temporarily disabling feelings are perfectly normal, e.g. in the face of the death of a loved one. And anyone may have feelings unwarranted by the objective situation, if he has been misinformed and thus came to hold false beliefs. By contrast, the mismatch between the subject's feelings and his own reflective beliefs

always constitutes a disorder – to which, ultimately, autonomous cognitive habits may give rise.

To cure an emotional disorder is to put an end to such pathological feelings. The 'cognitive approach' is to do this by identifying and 'modifying', i.e. weakening or breaking, the relevant cognitive habits. Beck found that in many cases this could be achieved simply by eliciting and refuting, again and again, autonomous thoughts that manifest the habit: this 'purely cognitive' exercise proved enough whenever the habit was not sustained by too powerful a psychological motivation. We shall see that the problems Wittgenstein tries to cope with in sections 138–97 of the *Investigations* are in some pertinent ways similar to the problems thus addressed by cognitive therapy: they arise in the context of autonomous thoughts, from autonomous cognitive habits, which Wittgenstein attacks much in line with the 'cognitive approach'.

3 Wittgenstein's autonomous thoughts

The discussion of *reading* revolves around the theme that it is a 'special conscious activity of mind' (156): that a conscious act constitutes the criterion for making the difference between actual reading and behaviour that merely looks like it (159), that real reading involves a highly characteristic experience (165), that when reading one feels a kind of influence of the letters (169), just as one has a particular experience of being guided by the original when transcribing a text (173) or copying a doodle (175). The sentences that express these different variations on the theme are prefaced by, or contain, odd riders that have, by and large, been ignored: '... we shall be inclined to say ...' (156e), 'So we should like to say ...' (156g), '... we are tempted to say: ...' (159a1), '– we should like to say –' (165a1), 'Here I should like to say ...' (165a9, 173b11), 'I should like to say ...' (169c2, 176a4, 177a1), '... and then you would like to say ...' (174a3), where 'should like to say' and 'would like to say' are of course translations of the same German expression: 'möchte ich/möchtest du/sagen'.

In line with English – though not German! – usage, Wittgenstein might be taken to use 'I should like to say that ...' quite simply to advance the claim that ... I will now argue that, first, it is indeed Wittgenstein who is speaking here in his own voice: he uses the 'I' to refer to himself. But, second, he is not at these points making the claim that ... First, while the general theme is natural enough (though not philosophically prominent), some of its variations are put by means of rather idiosyncratic formulations, and all of them are formulated in the context of quite *outré* behaviour which Wittgenstein describes in the first person: who would want to say, 'I experience the because' (176, 177)? Wittgenstein 'would like to say' this 'when I reflect on what I experience in such a case' as the one he considered a moment before (177), namely, when he is copying an arbitrary doodle, to pin down the 'experience of being guided' by the original doodle (175) – something no

other philosopher is known to have engaged in. And this latter point also holds of his less idiosyncratic inclinations: he would 'like to say: when I read I feel a kind of influence of the letters working on me – but I feel no influence from that series of arbitrary flourishes on what I say' (169c2) after having actually conducted the 'experiment' (169c8) of drawing a line of arbitrary flourishes and comparing the experience he has when reading a sentence with the one he has when looking along the line (169a2–3). There is no evidence to support the very unkind suggestion that Wittgenstein is setting up a remarkably unrealistic straw-man under the guise of a fictitious 'I'. Hence, we should assume that he here uses the first-person pronoun to refer to himself, and is sincerely stating what he 'would like to say' in the particular situations he describes.

But, second, in writing that he would like to say that . . . , he is not advancing the claim that In the first person, he tells us what he 'would like to say' mostly *after* having made the experiment he finds to refute the idea in question (169c2, 176a4, 177a1, following the experiments of 169 and 175). He then calls the idea a 'fiction' (166a2) or 'imagination' (170c1), and seeks to render the thought intelligible as a (misguided) interpretation of a salient feature of the experience had during the experiment at issue (170a, cf. 177), an interpretation that strikes him as particularly attractive 'when I say "guidance", "influence", and other such words to myself. "For surely", I tell myself, "I was being guided" – Only then does the idea of that ethereal, intangible influence arise' (175b7–9). Outside such reflection on those experiments, the thought 'would never have occurred' to anyone (170a1). So, far from endorsing them, Wittgenstein is, here, merely expressing ideas that occurred to him and struck him as attractive when describing a particular experiment he conducted, in particular terms (177a2), things he in this sense 'would like to say' – but which he rejects firmly enough to raise the question of how such things could occur to him, in the first place. (In 140a2, he wonders apropos another such thought: 'How could I think that?') Which question he answers for the ideas that instantiate the main theme on reading by conceptualising them as fictions conjured up under the influence of particular language.

Also, the other riders mentioned at the outset are used to express such ideas. In discussing the different things he 'would like to say', Wittgenstein is readily switching between the first persons singular and plural: what 'I should like to say' (165a8–9) renders more precise what 'we should like to say' (165a1–2); the next variation on the general theme, again: what 'I would like to say' (169c2), is a thought that 'would never have occurred to us . . . if we had not . . .' (170a1), spelled out with the words 'We imagine . . .' (170c1), but enlarged on with the words, 'For when I speak of . . . that is really meant to imply . . .' (170c2). Hence, also, in saying what 'we should like to say' Wittgenstein is, here, expressing ideas that, at certain points, struck him as attractive, even though he reflectively rejected them.

Very occasionally, Wittgenstein uses the related rider 'we are inclined to say' in making moves of a different kind, namely, in advancing claims usually associated with ordinary language analysis: in section I60, e.g. Wittgenstein imagines a situation (under the influence of some poison, someone systematically transforms into sounds signs not belonging to any existing alphabet) and ventures a hypothesis about how most competent speakers would describe it: 'here we should probably be inclined to say he was making up an alphabet for himself *ad hoc* and then reading accordingly' (my translation). The first articulation of the main theme on reading, though, is in a different boat: 'If we think of [one particular] sort of reading, the reading of a beginner [that was previously described], and ask ourselves what *reading* consists in, we shall be inclined to say: it is a special conscious activity of mind' (156e). 'We' are said to have this inclination when thinking of a quite specific situation (rather than: all our lives or 'whenever we philosophise'). But it is not a preference for one description of that situation (say, 'The pupil really reads') over another. It is an inclination to answer not the 'linguistic' question, 'How (as "reading" or "guessing", etc.) is the activity imagined or considered to be described?', but the entirely different and peculiarly philosophical question: 'What does *reading* consist in?'. Thus, sentences containing the rider 'we are inclined to say', just like those containing 'we should like' or 'are tempted to say', are frequently best regarded as expressing exactly what they purport to express: an *inclination* or *temptation* or *desire* to say certain things that Wittgenstein has and might share with kindred spirits – mostly things he, at the same time, reflectively refuses to endorse.

The riders mentioned at the outset mark only some of the inclinations and desires Wittgenstein attends to. Others are reported through – not all but many of the – remarks commonly attributed to a shadowy 'interlocutor' who is sometimes being addressed in the second person, if his remarks are not simply enclosed in double quotes. Wittgenstein, however, usually uses the 'you' to address himself, as one sometimes does in soliloquy ('But now I ask myself: What are you doing?' 173b8), and explicitly acknowledges as his own, crucial mistakes and wild ideas of the 'interlocutor' he addresses. Among these are instantiations of precisely the two core ideas we shall see to jointly raise the main problem of 138–97. The first surfaces in the mistake that 'we should like to express by saying: I should have thought the picture [that came before my mind when I understood "cube"] forced a particular use on me', which he acknowledges as 'my mistake' (140a1) after it has been made by 'you' in saying "It's quite simple . . ." (139d4–5). The second and related idea is first exposed when Wittgenstein turns from acts of understanding to acts of meaning: 'your idea was that that act of meaning the order had, in its own way, already traversed all those steps . . . Thus you were inclined to use such expressions as: "The steps are really already taken, even before I take them in writing or orally or in thought"' (188). Subsequently, Wittgenstein explicitly treats this idea and inclination as his own: "All the steps are really already taken"

means: I no longer have any choice. The rule, once stamped with a particular meaning, traces the lines along which it is to be followed through the whole of space. – But . . . my description only made sense if it was to be understood symbolically. – I [!] should have said: *This is how it strikes me* [!]’ (219).

Wittgenstein also reports such inclinations in double quotes that are not accompanied by an address in the second person, as the evolution of a related remark from its manuscript source shows:

“To us, a series only has *one face!*” – Certainly; but which? Surely the algebraic one, and that of a segment of its expansion. Or does it also have another one? – “But everything already lies in this!” Well, what more do you [!] want. That happens to be the exclamation that this situation brings about. – And it is now a different question: why I [!] am inclined to say precisely this. – For it // that // doesn’t belong to the application of the rule.

(MS 124, 184; my translation)

In the course of this soliloquy, Wittgenstein does address himself in the second person. In the section derived from this, however, no such pronoun occurs; what he himself is inclined to say now simply appears in double quotes:

“To us, a series only has *one face!*” – Certainly; but which? Surely the algebraic one, and that of a segment of its development. Or does it also have another one? – “But everything already lies in this!” – But that is not an observation about the segment of the series, or about anything that we notice in it; it rather gives expression to the fact that we only look at the mouth of the rule and *do* [*sic*], and do not appeal to anything else for guidance.

(228, my translation)

[The last sentence gives one sort of explanation of ‘why I am inclined to say precisely this’ – note the idiosyncratic phrasing, as awkward in German as it is in English, which Anscombe smoothed down.]

Wittgenstein came to attach considerable importance to this exercise of attending to inclinations to talk nonsense, to ideas that in certain situations strike one as plausible and attractive, even though one reflectively rejects them all along. Thus he exhorted himself about a year after finishing his work on the *Investigations*: ‘Don’t, for Heaven’s sake, be afraid of talking nonsense! Only don’t fail to pay attention [literal translation: to listen] to your nonsense’ (MS 134, 20; 5.3.1947, CV 64). Already, sixteen years earlier, he had insisted upon the importance of giving so ‘characteristic’ an expression to wrong lines of thought that it is being accepted as ‘the true expression of [one’s] feeling’ (FF 108b, derived from BT 410). By the time of the

Zwischenfassung, he had come to conceive of this as the task of giving ‘a psychologically accurate exposition of the temptation to say this-or-that’ (ZF 216, early version of section 254). In the light of this remark, some of the relevant manuscript material, like the extract just quoted (and its surround: MS 124, 184–5, quoted below), reads like records of deliberate efforts to attend to such ‘temptations’, to what one is inclined to say when relinquishing deliberate control over one’s thinking and allowing oneself to say ‘how things strike one’, no matter how nonsensical: thoughts that occur to Wittgenstein in the course of this exercise (rendered for the most part in double quotes) alternate with comments on them made in cold blood, and responses to these comments that occur to him spontaneously. Many of these thoughts then recur as variations on the same general theme, even once the general idea and the thoughts previously instantiating it have been reflectively rejected: take the things Wittgenstein ‘would like to say’ on reading and copying. Both parts of our first definition are thus satisfied: these and the other ideas we discussed in this section occurred to Wittgenstein as *autonomous thoughts*.

4 Wittgenstein’s cognitive habits

But why did he explicitly attend to them? Why should a philosopher bother to take note of things that may occur to him, when he reflectively rejects them or, indeed, coolly judges them to be not even faintly sensible? This and the next section build up to the two parts of an answer: Wittgenstein’s autonomous thoughts are part and parcel of two autonomous cognitive habits. And cognitive distortions in line with these two habits generate the puzzles from which Wittgenstein, here, seeks to rid himself.

Practically all the autonomous thoughts reported or emerging in sections 138–97 are instantiations of, or variations on, two general themes. One of them we already encountered in the discussion of reading. In more general terms, it can be characterised as concerning ‘context-sensitive activities’ like reading, taking down a dictation, or copying a text or figure: these activities are associated with characteristic forms of behaviour such as, e.g. looking at a text and following it line by line with one’s eyes. But it is conceivable that the characteristic behaviour is performed without qualifying as reading, copying, etc. Indeed, it actually happens that pupils recite a text they know by heart while looking at the text and even following it line by line with a finger, thus pretending to be reading it off when they do not. In such cases, we make the difference between true and merely apparent V-ing by reference to behaviour in other, actual or possible, situations: we consider, e.g. whether the child is able to recite the same text without consulting it, or to read out another text of comparable difficulty. But the moment they think about it in the abstract, many people find it plausible to say that this difference has to be made by some mental or inner event. Of course, Wittgenstein sincerely and reflectively repudiated this idea. But, even so, he has autonomous

thoughts according with this idea, not only on context-sensitive activities but also on understanding, in particular, on understanding of the system of a number-series. While of course no activity, this particular kind of understanding is similar to context-sensitive activities in that it goes with highly characteristic behaviour: continuing a segment of the series, uttering a formula, etc.; but, as in the case of reading etc., it is conceivable that a subject behaves in these ways and yet fails to understand. It is, thus, with respect to everything that strikes him as relevantly similar to context-sensitive activities, that Wittgenstein is tempted to say things in line with:

Schema A: When someone actually V-ies, (1) a mental event takes place that makes the difference between true and merely apparent V-ing, and (2) makes him act as appropriate, with more than merely causal inexorability.

The working of an inner mechanism serves here as a model of 'mental events' (156g). Schema and model jointly constitute what can aptly be called the '*mentalistic picture of context-sensitive activities*'.

The second theme concerns logical or 'grammatical' determination: we sometimes say that a conclusion is 'already contained in' the premises that logically entail it. Again, Wittgenstein is inclined to say something along these lines in a host of further cases. For example, the way in which we use a word is one criterion of how we understand it. Conversely, the assumption that I do understand that word correctly generally suggests that I will use it correctly, and would seem to logically entail this, in conjunction with various other assumptions (of sobriety, care, etc.). In a way, therefore, my understanding of the term seems to determine how I will use it, and to determine this with a 'grammatical' inexorability comparable to that of logic. In many such cases Wittgenstein feels inclined to say things in line with another assumption he never reflectively endorsed:

Schema B: When something determines something else with stronger than merely causal inexorability, the former somehow already contains or anticipates the latter, i.e. brings it into some sort of existence before it actually comes about.

The workings of an 'ideally rigid' ethereal mechanism or machine, then, are the model for 'more than merely causal inexorability' (193). This schema and model jointly make up what can fittingly be called the '*actualistic picture of logical determination*'.

The schemata and the models that go with them bias both Wittgenstein's automatic and his more controlled thinking, issuing both in the autonomous thoughts that occur to him when 'listening to' his inclinations and temptations to talk 'nonsense', and in conclusions he may momentarily endorse.

These conclusions are manifestly reached through leaps of thought that tacitly presuppose assumptions instantiating those schemata. Many of those autonomous thoughts are analysed by Wittgenstein himself as misinterpretations informed by such assumptions. Let's now take note of the autonomous thoughts he reports and the leaps we can observe in his controlled reflections.

The textual markers discussed above allow us to identify the following reports of autonomous thoughts, in sections 138–97:

- (140a1) The picture we grasp when understanding a word forces us to apply it in a certain way, with logical compulsion. (Schema A1&2)
- (146a6–b2) Understanding (of a number series) is a state from which the correct development flows, like the derivation from an algebraic expression. (A1&2)
- (147a2, b1) When I say I understand the rule, I am not saying so because I have found out that up to now I have applied the formula in such-and-such a way. I know the application apart from remembering actual applications to particular numbers. (Correct as it stands, but in 147a1 interpreted in line with A1)
- (156e) Reading is a conscious activity of mind. (A1)
- (156g) Two different processes, or mechanisms, are going on in the proficient reader and the beginner (who is pretending), and this difference in what is going on in them distinguishes reading from not reading. (A1, model A)
- (159a1) The one real criterion for anybody's reading is the conscious act of reading, the act of reading the sounds off from the letters. (A1)
- (165a1–2) Reading is a quite particular process, with something special and highly characteristic going on. (A1)
- (165a9) The words I utter come in a special way. (A1)
- (169a1) When we read, don't we feel the word-shapes somehow causing our utterance? (A1, model A)
- (169c3) When I read I feel a kind of influence of the letters working on me – but I feel no influence from that series of arbitrary squiggles on what I say. (A1, model A)
- (170c1) A feeling enables us to perceive as it were a connecting mechanism between the look of the word and the sound that we utter. (A1, model A)
- (173a1) But being guided surely is a particular experience. (cf. 165a1–2. A1)
- (173b11) The experience of being guided is something more inward, more essential. (A1)

- (174a3) The experience of deliberation is a particular inner experience. (Extension of A1 from 'context-sensitive' verbs to adverbs qualifying 'conscious activities')
- (176a4, 177a1) When being led by a figure I copy, I experienced the because. (A1, model A)
- (184a5) When I suddenly remember a tune, I have the distinct feeling as if it were there. (B)
- (188a) That act of meaning the order had in its own way already traversed all those steps: when you meant it your mind as it were flew ahead and took all the steps before you physically arrived at this or that one. (B)
- (188b) The steps are really already taken, even before I take them in writing or orally or in thought. (B)
- (191a1, 197a1) It is as if we could grasp the whole use of the word in a flash. (Where the misplaced caution - 'as if' - is due to B, see next section)
- (193a2) The action of a machine seems to be there in it from the start. (Consequent B)
- (193c1) The future movements of the machine are, in their definiteness, like objects which are lying in a drawer and which we then take out. (Consequent B)
- (194b5, 7) Possibility is something which is like reality. Possibility is something very near reality.
- (194b9, 10) The possibility of the movement stands in a unique relation to the movement itself; closer than that of a picture to its subject. It is not an empirical fact that this possibility is the possibility of precisely this movement. (Antecedent B)
- (195a1) What I do in grasping determines the future use of the word not causally and as a matter of experience; rather, in some queer way the use itself is in some sense present. (B)
- (197a5) The future development must in some way already be present in the act of grasping and yet isn't present. (Consequent B)

The thoughts clearly reported as autonomous thoughts thus mostly instantiate schemata (B) and (A1). Thoughts instantiating (A2) are discussed in sections 139c-146 and 181-3: when the picture of a cube comes before my mind in understanding the word 'cube', I *must* go on to apply the word precisely to cubes (140a1-2); we *must* go on thus in case the picture is supplemented by a projection schema (141a1-4). Once Wittgenstein has made it clear that the picture might as well exist as a drawing in front of the subject, rather than in his imagination (141b3), he seeks to dislodge the idea

that when told to copy the picture '0, 1, 2, 3, 4, 5, 6, 7, 8, 9' on the board a pupil will *inevitably* put down this number-series (143), and the thought that the pupil *must* continue the series of natural numbers correctly when given an explanation employing some such picture as (145a):

_0, _1, _2, _3, _4, _5, _6, _7, _8, _9
10, 11, 12, 13, 14, 15, 16, 17, 18, 19

But he still feels inclined to voice the idea that the understanding of a number-series is the source from which the correct use flows like a series (inexorably) derived from an algebraic formula, a (mental) state that *forces* the subject to continue the series correctly (146a6-b2). Variations on this theme recur once Wittgenstein has clarified that we use the exclamation 'Now I know how to go on!' rather like a signal, the correctness of whose use depends upon the circumstances: when the right formula occurs to us, e.g. we may say this if, and only if, we have learnt algebra and have already used formulas of the kind in question before (179-80). This observation gives rise to the thoughts: 'When it is under the right circumstances that a subject says, "Now I know how to go on", then he *must* subsequently be able to continue the series correctly' (cf. 181), and 'There is a totality of conditions such that someone *could not but* go on correctly if all of them were fulfilled' (cf. 183). While he explicitly reports only two of them (140a1-2, 146a6-b2) as autonomous thoughts, Wittgenstein repeatedly tells himself to 'be on [his] guard against thinking' such things (183a5), wants to get himself to acknowledge that other responses than the supposedly inevitable ones might occur as well (144a6), and finds it necessary to remind himself of these other possibilities, even though they really are obvious (145a6). The ideas in line with (A2) thus manifestly struck him as plausible, even forceful, while he was throughout careful not to endorse them. I therefore propose to regard them, too, as autonomous thoughts of his.

The schemata and models identified do not merely surface in such thoughts. Both bias systematically all of Wittgenstein's thinking, including his controlled reflections: as we will show in detail later, those schemata and models jointly distort his interpretation of thoughts and utterances, experiences and events. And, as we presently find out, Wittgenstein regularly draws inferences in accord with the schemata. With systematic inadvertence, assumptions instantiating them wholly or in part are being tacitly presupposed in leaps of thought that lead him precisely into the three principal puzzles of 138-97: into the difficulties about supposedly sudden understanding raised in sections 138-9, into the muddle about strangely elusive mental events involved in genuinely sudden understanding that he develops in sections 151-3, and into the full-blown puzzle about, again, supposedly sudden understanding that he treats in 191-7.

Let's consider first the least complex second of these problems. Wittgenstein develops it from the first-person perspective, changing between the first persons singular and plural, to sum up: 'I am in a muddle' (153a7) – clearly he himself felt muddled and confused. On the face of it, the muddle is generated quite simply by an attempt Wittgenstein makes to identify a mental event that will fit the bill of assumption (A): a mental event that (1) occurs whenever someone suddenly understands the system of a number-series, and (2) inexorably makes the subject act as appropriate. For the mental events that readily come to mind as accompaniments of the sincere exclamation 'Now I understand!' (which Wittgenstein goes through in 151) need not go with the ability to continue the series correctly (as he seems to observe in 152). Hence, the true mental event or process of understanding seems to him hidden behind these mere accompaniments (153a1). And this concealment seems strange because he thought he exclaimed 'Now I understand' *because* he understood (153a5), which he seems tempted to interpret as meaning: because he grew aware of the relevant mental process.

But, of course, Wittgenstein does not endorse any such assumption as (A). From the start (151b4), he equates the pertinent understanding with an *ability*; in the end (154c1, d), he explicitly rejects the opposing idea (A1) of understanding as a mental process; and, all along, he countenances only what 'remains' of (A2) after 'cleansing' it of this idea. The way an algebraic expression is applied is one criterion of how it is understood (146b6), so that the attribution of its correct understanding 'inexorably' implies that of the ability to correctly continue the series. So, of course, Wittgenstein does not deliberately look for a mental event that satisfies the very bill (A) makes out. Rather, he is led into the muddle by reflection that crucially involves three presumably inadvertent leaps of thought in line with precisely the first part of schema (A).

The first leap is from 151b5 to 151b7. Considering the case of a pupil who suddenly understands the system of a number-series and is, thus, able to continue the segment so far developed on the board, Wittgenstein asks: 'So this capacity, this understanding, is something that makes its appearance in a moment. So let us try and see what it is that makes its appearance here' (151b4–5). As he explicitly equates understanding of the system with the capacity to continue the series, one would expect him to answer the question 'What is it that makes its appearance here?' by some more detailed or otherwise informative description of that capacity that is what makes its appearance here. But, instead, Wittgenstein jumps to the question 'What happened here?' (151b7), which is a question about a – presumably mental or inner – event.

The second leap, from the mental events described in reply, in 151b–d, to the question 152a, involves a weaker version of the same presupposition (A1): 'Are the processes which I have identified the understanding?'. Even when a negative answer is expected, the question presupposes that the

understanding might very well be some such mental event or process as the ones he described. Wittgenstein tacitly recasts the question as that of whether attributions of those processes and of understanding mean the same, and gives a negative answer because 'it is perfectly conceivable that the formula should occur to him [etc.] and that he should nevertheless not understand' (152b1–2). Presumably, he has in mind here the possibility that the subject might fail to continue the series correctly. The conflation of the two possibilities is, of course, unproblematic, as it is in line with both (A2) *and* his explicit view on criteria of algebraic understanding. Not so Wittgenstein's final leap of thought, from 152b2–4 to 153a1, as analysed by himself: from the fact that 'Now I understand the system' does not mean the same as 'The formula occurs to me', he jumps to the conclusion that the former must describe a process occurring behind that of saying (*sc.*: or thinking of) the formula (154a). Which move relies on the 'linguistic' version of (A1): attributions of understanding describe mental processes. The three vicious leaps of thought forcefully illustrate how the habit of inadvertently drawing inferences in line with assumptions he reflectively rejects may lead a philosopher into a muddle he has to make some efforts to dissolve to his own satisfaction. Table 5.1, below (in section 6), gives an overview over these and other leaps of thought Wittgenstein makes in sections 138–97, and the perplexities to which they lead.

5 The main problem of sections 138–97

Controlled reflection marred by cognitive distortions in line with the mentalistic picture (A) thus gets Wittgenstein into the muddle of sections 151–3. The main puzzle, first mooted in sections 138–9 and fully spelled out in sections 191–7, is a more complex affair, the result of controlled reflection *and* automatic thought in line with *both* the mentalistic picture (A) and the actualistic picture (B) – and, indeed, with the meaning-body conception that frequently goes with the Augustinian picture of language. To clarify the nature of the final problem, we will first consider section 197. To unravel the lines of thought that led Wittgenstein into it, we shall then move back to sections 138–9.

The discussion of supposedly sudden understanding, the main theme of sections 138–97, culminates in the discussion of the impression that 'it is as if we could grasp the whole use of a word in a flash' (191a1, 197a1), namely, when we hear or say the word with understanding (138a4, cf. 197a6). This remark articulates the feeling that something 'astonishing' and 'queer' is happening when we understand a word (cf. 197a4). Indeed, that we should be able to grasp the *whole* use in a *single* instant (cf. 139b3–4) seems *so* astonishing and queer that the thought stops short of suggesting that such a thing might actually happen: 'It is *as if* we could grasp . . .'. This sense of wonder is of the peculiar kind Plato took for the starting-point of

philosophising: a feeling of amazement or wonder in the face of an entirely familiar and pedestrian phenomenon – like understanding a word we hear or say.

From where does this feeling come? According to Wittgenstein's *own* analysis, 'It becomes queer when we are led to think that the future development must in some way already be present in the act of grasping and yet isn't present' (197a5, my translation) – a thought that is, indeed, apt to induce intellectual vertigo. In the *Frühfassung*, the later sections 191–7 form part of the discussion of the 'hardness of the logical [and mathematical] "must"' (FF 349–57). This context renders it less than surprising that the present thought is yet another variation on the theme that 'everything already lies' in what comes before our mind when we consider the segment of a number-series with understanding. This is the theme of various things Wittgenstein explicitly reports he is inclined to say about that 'hardness', without ever explicitly countenancing them. Clearly, then, Wittgenstein in 197a5 reports an autonomous thought he had himself. This autonomous thought gave rise, in him, to the feeling of amazement and wonder he adverts to in 197a4. The opening sentence articulating this feeling (197a1) sums up how, in the grip of that thought, things struck him (cf. 219).

But why should Wittgenstein (or anyone else, for that matter) be led to think that 'the future development [or use] must in some way already be present in the act of grasping'? He answers: 'For we say that there isn't any doubt that we understand the word, and on the other hand its meaning lies in its use' (197a6). It is a platitude that (unless we travel abroad) we mostly understand the words we hear or say. And while Wittgenstein was, in general, careful not to endorse answers to the dubiously meaningful question 'What is the meaning of a word?' (as evidenced, e.g. by the – in German – cautious formulation in section 43, the mere conditional in 138, and the double quotes in 560), he arguably did countenance the use of 'The meaning of a word lies in its use' as a slogan, to sum up facts like these: when we ask what the meaning of a certain word is we want to know how it is used or how we should use it; when we explain its meaning we explain the manner in which it is to be used; we often ask 'What do you mean by that word?' when someone uses it in an unusual way and we want to know how *he* uses it. The slogan should, hence, be read as: 'The meaning lies in the word's way or *manner* of use'. But this does not even faintly suggest that the *future* use has to be 'present in the act of grasping': quite obviously, I do not know how I *will* use a word when I know how it *is to be* used; for all I know, I might get drunk or original. Second, Wittgenstein appears to have in mind the understanding we can be credited with when hearing or saying a word we understand (138a4). On his considered view, this 'understanding' consists in nothing over and above the general competence of using the word: the possession of this general competence which may manifest itself in various different situations, rather than any mental act or process, makes the relevant difference between someone who

hears or utters a word with understanding and someone who hears or parrots it without – this, at any rate, is the principal upshot of sections 138–83. Properly interpreted in line with Wittgenstein's considered views, the platitude and slogan (of 197a6) thus do not even remotely suggest that an 'act of grasping' occurs whenever we hear or say a word we are familiar with, let alone that its 'future use' or 'development' (195a1, 197a1) should be in any way present when we know how it is to be used.

To sum up: Wittgenstein puts his feeling of amazement and wonder down to what we learned to conceptualise as an autonomous thought (197a5). And this thought is entirely unreasonable in the light of his two pertinent considered beliefs (197a6). So is his peculiarly philosophical sense of wonder. This feeling, thus, is in one crucial respect analogous to the feeling of anxiety experienced by Beck's patient: it is intelligible in the light of an explicitly reported autonomous thought but unreasonable in the light of the beliefs he reflectively holds at the same time.

But, reasonable or not, Wittgenstein actually did jump from the platitude and slogan he identified as his point of departure to the ideas that are apt to induce his philosophical sense of wonder. He did so by two leaps of thought that implicitly rely on precisely the two assumptions he is most intent on unsettling in the preceding part of the *Investigations*: on the schemata constitutive, respectively, of the mentalistic picture of understanding and of the meaning-body conception that frequently goes with the Augustinian picture of language. He makes both leaps in section 138. The first is this: 'We *understand* the meaning of a word when we hear or say it; we *grasp* it in a flash' (138a4). This move is odd: we only say that we 'grasp the meaning of a word in a flash' when we pick it up in a sufficiently informative context (rather than looking it up in a dictionary or being told). We then say something about how we first came to understand the word. In advance of some innovative explanation of meaning, it is, therefore, not even meaningful to say that we 'grasp a word's meaning in a flash' when we hear or say a word we are already familiar with (though we can, of course, say that we grasped in a flash the import or implications of what was said by an utterance of the complete sentence, or the allusion made through the particular choice of words). In making this move without any explanation whatsoever, Wittgenstein is relying on an instance of schema (A1):

When someone says or hears a word with understanding, a mental event takes place that makes the difference between this and merely apparent understanding (merely parroting the word or listening with an unwarrantedly knowing smile).

It is then tempting and natural to describe this event in familiar terms usually reserved for more specific occasions: 'The subject grasps the word's meaning'. In particular, in conjunction with model (A) of a mental event as

the working of an inner mechanism, this formulation then suggests that a certain link or relation is being established between the subject and the word's meaning. Which renders it natural for Wittgenstein to slip from saying that 'we understand a word' into saying that 'we understand the meaning of a word' (138a4), which is as unidiomatic in German as it is in English (you know, or are familiar with, either a word or its meaning, but are said to 'understand' only the former). Thus, (A1) and the model it goes with first influence Wittgenstein's formulation of a platitude and then have him jump from it to a piece of nonsense, namely, from

- (1) 'When we hear or say a word we (usually) understand it', to
- (2) 'When we hear or say a word we (usually) grasp its meaning'.

The second leap is constituted by an odd interpretation of his own slogan that 'the meaning is the *use* we make of the word' (138a3), which we saw to amount to: 'The meaning is the way or *manner* in which the word is used'. In raising the objection that 'what we grasp in this way is surely something different from the "use" which is extended in time' (138a4), Wittgenstein interprets it differently: of course, also, a manner of use can be said to have a temporal extension, as words acquire a certain use and lose it again (rather quickly, e.g. in the case of youth language). But there is no obvious difficulty in the idea that we come to understand in a flash one such way in which the word may be used. As Wittgenstein does not bother to explain any less than obvious difficulty, he must mean something different by 'the use which is extended in time': he interprets it in line with an assumption articulating the 'meaning-body conception' that informs the discussion of 136b–138a3, namely,

The meaning of a word is an object – with a temporal, if not a spatio-temporal extension.

In section 138, he thus interprets the phrase as meaning 'the several instances of the word's use that are spread out over time', before focusing, in sections 191–7, on the *future* instances of use (195a1, 197a5). He thus manifestly jumps from

- (3) 'The meaning is the *use* we make of the word', to
- (4) 'The meaning is the sum of the instances of the word's use'.

While we saw three paragraphs up that (1) and (3) don't even suggest a puzzle, (2) and (4) do imply a highly perplexing conclusion:

- (5) When we hear or say a word, we (usually) grasp all of the instances of its use.

And this, Wittgenstein spontaneously protests, cannot be (138a4).

He then considers the question by means of an example: 'When someone says the word "cube" to me . . . I know what it means. But can the whole use of the word come before my mind, when I *understand* it in this way?' (139a). He finds it so puzzling that the whole use of the word (spread out over such a long period of time!) should be present to us in a single instant (cf. 139b3–4), that he immediately gives a negative answer: 'No, what can come before my mind in an instant, what I can grasp in a flash, is not the whole use or *meaning* of the word, but something that *determines* this meaning.' This implicit answer emerges from the explicit continuation: 'Well, but on the other hand isn't the meaning of the word also [!] determined by this use? And can't these two ways of determining the word's meaning [through what is being grasped, on the one hand, and through the use made of the word over time, on the other] conflict?' (139b1–2). That is: faced with the consequence (5), Wittgenstein suddenly finds the idea of 'grasping a word's meaning' so puzzling that he retracts it, and moves from (2) to the idea that when we hear or say a word, we grasp something that *determines* the meaning, which he again equates with the uses or applications made of the word. That we should actually be able to grasp the meaning or use itself, seems too mind-boggling to be true. Already at this point, he is presumably seized by a sense of wonder that can be put into the words: "'It is *as if* we could grasp the whole use of a word in a flash'" (191a1, my italics). Precisely to escape it, he moves to a new thought, that instantiates schema (A2):

- (6) When I hear or say a word with understanding, something comes before my mind and determines the use I make of the word, with logical rather than merely psychological compulsion.

Thus, the 'mistake' he diagnoses as his own: 'I should have thought the picture [I supposed to come to my mind when I hear the word "cube"] forced a particular use on me' (140a1–2), namely, forced him to apply the word precisely to cubes, with logical compulsion rather than the 'merely psychological compulsion' that, he is subsequently inclined to say, is all there is to it (140a4–5).

But the move from the idea that I grasp the meaning or use of a word when I hear it to the idea that I then grasp something that determines its meaning or use brings him no lasting relief from puzzlement. An autonomous thought, reported in section 195, infers from the new assumption (6) a conclusion that is slightly different from but no less puzzling than (5), namely, the conclusion that the *future* instances of use must somehow already be present in the act of grasping: "'But I don't mean that what I do now (in grasping) determines the future use *causally* and as a matter of experience, but that in a *queer* way, the use itself is in some sense present'" (195a1, my translation). The first, explicit, premise is that (6) what comes before my mind when I understand a word, what I then grasp, determines the future use with not

merely causal inexorability – but with logical compulsion. The second, tacit, premise is uncovered by one of three diagnostic remarks that separate the counterparts of sections 195 and 197 in the *Frühfassung* (FF 355). It is a strong version of schema B:

When something determines something else with logical compulsion, the latter somehow already is the former.

These two premises jointly imply that:

- (7) When I hear or say a word with understanding, I grasp something that somehow already is the future use of the word

– so that ‘the future use is in some sense present’ in the act of grasping (as 195a1 concludes). Which is the thought to which Wittgenstein puts down his sense of wonder (in 197a5; Anscombe’s translation of 195a1 and 197a5 obscures the subtlety by turning what is in German an ‘act of grasping’, *tout court*, into an ‘act of grasping a sense’ or ‘the use’). The slip in this sentence (197a5), which speaks of ‘the future development’ instead of ‘the future use’, betrays the influence of his autonomous thought that ‘everything already lies’ in what we grasp when understanding a number-series, which may serve to illuminate the related present conclusion:

[When I consider the segment with understanding] I believe to see a design drawn very fine in a bit of a series [A1], which only stands in need of “and so on” to reach to infinity [B].

“I see a distinctive character in it.” – Well, presumably something that corresponds to the algebraic expression. – “Yes, only nothing written, but positively something ethereal.” – What a queer picture. – “Something that is not the algebraic expression, something for which this is only the *expression!*” [A1, model B]

(MS 124, 184–5, kept as Z 276)

[Sequel to remark “But everything already lies in this!” quoted in section 3, entire passage source of PI 228–9.]

In other words, when I consider the segment of a series (and understand its system), I discern in it, or grasp, something ethereal, like a design with an ‘and so on’ attached, that somehow already comprises the whole expansion, including all of its future development.

The influence of the actualistic picture of logical determination also accounts, I think, for the most glaring inconsistency of sections 138–9. First, Wittgenstein implies that the meaning of a word is its use (138a3), and then states clearly that the use determines the meaning (139b1). But, one would like to object, a thing cannot be determined by itself! However, the

determination in question is ‘more than merely causal’, logical, determination, so that, by (B), the use ‘somehow already is’ the meaning. The actualistic picture thus affected Wittgenstein’s thinking not only in sections 191–7 but also in 138–9, when first mooting the problem.

To sum up, two distinct lines of thought lead Wittgenstein to a thought that makes him feel puzzled, even though he does not endorse it: to the perplexing idea that when we hear or say a word with understanding there occurs an act of grasping in which all the (future) instances of the word’s use are somehow present. Wittgenstein rehearses the first line of thought in controlled reasoning, the second occurs in automatic thought. The first proceeds *via* the notion that whenever I hear or say a word with understanding, I grasp its meaning. The second starts out, instead, from the idea that I then grasp something that determines its meaning. Wittgenstein moves from one notion to the other, precisely to avoid the perplexing conclusion. But he is confronted with it, either way. In his deliberate reflection, two leaps of thought in line with the mentalistic picture of understanding (A) and the meaning-body picture, respectively, lead him from platitudes to perplexity. In his automatic thinking, a leap of thought in line with the actualistic picture (B) takes him there from a slightly different point of departure. The influence of all three pictures is evident in both the reflection and the automatic thought. The main perplexity of sections 138–97, a peculiarly philosophical sense of wonder in the face of a pedestrian phenomenon, thus results from the interplay of different autonomous habits of thought. These findings will allow us to make good sense, first, of Wittgenstein’s general aims and, then, of the way in which he pursues them in sections 138–97, i.e. of his ‘therapeutic’ approach.

6 Wittgenstein’s declared aims

In the year he engaged upon the second and final major round of work on the *Investigations*, Wittgenstein concisely characterised the philosophical task he set himself: ‘The philosopher is someone who has to cure in himself many diseases of the intellect [*Krankheiten des Verstandes*], before he can arrive at the notions of common sense’ (MS 127, 76r: 1944; CV 50). This has two clear implications. First, at least the immediate aim of ‘the philosopher’, i.e. Wittgenstein, is simply to return to notions of common sense from which he strayed. Second, diseases of the intellect prevent him from apprehending these notions properly. To achieve his aim, he therefore has to address the therapeutic task of curing those diseases in himself.

Both the, at first sight ridiculously modest, aim and a motivation that lend it a point are prefigured already in the earliest methodological remarks (from 1931) published in the *Investigations*: ‘If one tried to advance *theses* in philosophy, it would never be possible to debate them because everyone would agree to them. The aspects of things that are most important for us [namely, when doing philosophy] are hidden through their simplicity and familiarity’

(128–9a1). The exercise of getting the humdrum back into view is worthwhile because it puts an end to self-made mystery: ‘The results of philosophy are the uncovering of one or another piece of plain nonsense, and bumps that the intellect has got by running its head up against the limits of language. These bumps make us see the value of the discovery’ (119, my translation; Anscombe completely distorts the meaning by imagining an ‘of’ before the first ‘bumps’; note the comma before ‘&’ in the *Urfassung*). That is, philosophical activity has got two kinds of results. Some such activity results in ‘bumps’ – take, e.g. the reflections of sections 138–9 and 151–3 that lead Wittgenstein from familiar facts and pedestrian observations into puzzlement and confusion, the ‘bumps’. Other philosophical work results in the discovery of plain ‘nonsense’ in such reflections. The ‘bumps’ lend this discovery what value Wittgenstein can discern in it, i.e. he uncovers ‘nonsense’, quite simply, so as to dispel the puzzlement and confusion engendered by the other, vicious, kind of philosophical reflection. His main motivation is the desire to escape from bogus mystery. To this end, he seeks to return to the notions of common sense from which philosophical reflection led him into such mystery.

Also the diagnosis is prefigured in the earlier remarks: ‘Philosophical problems . . . are solved . . . *against* a drive to misunderstand [the workings of our language]’ (109a8, my translation), and at least many of them arise from such and related misunderstanding (111a1). The intellectual *drive* thus also generates many of those puzzles and confusions, in the first place. Note the implications of the strong term ‘*Trieb*’ (which Anscombe renders as ‘urge’): you find a drive in yourself, like the innate drives for sex and food; even if not innate, you experience it as a *datum*. And it asserts itself in certain situations while lying dormant in others; when asserting itself, it may drive you with great force, controllable only with a considerable effort (typically absent in the case of the intellectual drive at issue). The drive to misunderstand can be dysfunctional, in this technical sense. In particular (though not exclusively) if the person affected is a philosopher, it will ‘interfere with his occupational functioning’. Non-somatic pathology is commonly regarded as a matter of person-relative dysfunctionality. A non-somatic condition is ‘pathological’ if it is ‘disabling’, i.e. ‘interferes with the occupational or social functioning’ of the particular person in it (where one and the same urge, say, may interfere with the occupational functioning of one person but not another who has a different occupation). In this well-established sense, the present drive is pathological in a philosopher. Experienced as given, occasionally acute and, in philosophers, pathological, this intellectual drive can, indeed, be aptly termed a ‘disease of the intellect’. This diagnosis and the aim and motivation previously indicated jointly provide the rationale of the notoriously ill-understood project of a philosophical therapy. They have struck many commentators as over the top (if not below the belt). Now that we have conceptualised the main problems of sections 138–97 as the result of

autonomous cognitive habits, we are in a position to make good sense of all three.

At any rate some of the ‘drives to misunderstand’ he diagnosed, arguably in himself, are such habits of thought. He is manifestly driven to place puzzling interpretations like (2) and (4) on pedestrian facts and innocuous slogans like (1) and (3) and, thus, to jump to perplexing conclusions like (5) and (7). These leaps are part and parcel of comprehensive cognitive habits. Such habits may be ‘pathological’ also in a less pragmatic sense than the one we just encountered. In the case where the subject is not merely unaware of relying on assumptions that instantiate the defining schema of the habit but reflectively rejects all such assumptions, the habit is pathological also in a sense analogous to that applying to the feelings involved in emotional disorders (explained in section 2). This is true of Wittgenstein’s habits (A) and (B). Autonomous cognitive habits like these are his ‘diseases of the intellect’; each pathological habit is one such ‘disease’.

To see how the predicament thus diagnosed lends a point to Wittgenstein’s aim, we need to clarify the vicious effects of these habits: to become clear on the extent to which they bias his thinking, and on the relevance of the biased interpretations and leaps of thought that result. To this end, cognitive therapists employ the useful ‘ABC-schema’. ‘A’ stands for the ‘activating event’ or fact the distorted thoughts are thoughts about, under a straightforward and uncontroversial description the subject acknowledges as correct. ‘B’ stands for the distorted, mainly autonomous, thoughts at issue (often misleadingly called ‘beliefs’), as well as for the tacit assumptions the subject acknowledges. ‘C’, finally, stands for the ‘emotional and behavioural consequences’: the feelings, actions, and urges to act, with which the subject responds to the ‘activating event’, that are intelligible only in the light of the previous Bs. Table 5.1 gives such an overview over those of the puzzles and muddles of sections 138–97 that we have analysed or at least touched on, so far.

The findings summed up by Table 5.1 allow us to spell out in detail what Wittgenstein’s immediate aim comes down to, and to render his motivation intelligible. In all cases analysed, he leaps from, ultimately, a statement, finding, observation or question that is perfectly true, correct, or sensible – and, as often as not, rather pedestrian – to statements, conclusions or questions which are not. This happens in two ways: inadvertent misinterpretation and mindless inference. Both yield results that may be met with either of two attitudes. Wittgenstein thoughtlessly countenances some at first, while, throughout, he is careful not to endorse others. The latter are autonomous thoughts that capture ‘how things strike him’. Inadvertent misinterpretation occurs in cases (a), (b), (d), (e) and (f). Here, Wittgenstein simply mistakes the question or statement he starts out from (e.g. ‘I was guided’) for another (‘I had a characteristic experience of being guided’); he inadvertently interprets the former as amounting to the latter (and only subsequently recalls, e.g. that his ‘being guided’ consisted in drawing one line parallel to the other,

Table 5.1

A	B	C
(a) 138-9 (1) When we hear or say a word, we understand it (138a4) (3) The meaning is the use we make of the word (138a3)	(2) When we hear or say a word, we grasp its meaning (138a4) (4) The meaning is the sum of the instances of the word's use. Whence: (5) When we hear or say a word, we should grasp its whole use, extended in time – which cannot be! (138a4)	Sense of wonder articulated by: 'It is as if we could grasp the whole use of the word in a flash' (191/7a1)
(b) 151 Considering an example: upon finding a law for the number-series presented, B exclaims: 'Now I can go on!' (151b1-3), and asking: 'What is it that makes its appearance here?' (151b5)	Interpreted in the light of A1 as: 'What happened here?' (151b7)	Brainstorming to come up with various (mental) events that may occur in such a situation (151b-d)
(c) 152-3 Finding that all the events he came up with are compatible with failure to understand (152b2-4)	A1&2: 'There <i>must</i> be an event that guarantees correct continuation of the series' tacitly presupposed in leap to conclusion: the true mental event of understanding is hidden behind the coarser accompaniments that readily sprang to mind (153a1)	Urge to search and grasp this hidden mental event (153a1)
(d) 153 Observation that I said 'Now I understand' <i>because</i> I understood (153a5)	Interpreted in the light of A1 as: I said 'Now I understand' <i>because</i> I noticed the mental event	Feeling of confusion in the face of the previous conclusion that the event is hidden: 'I am in a muddle' (153a7)
(e) 153 'I say [the mental event of understanding] is hidden' (153a6) and, at the same time, that I want to look for it	The interpretation as analogous to 'I am looking for my key' rather than 'I am looking for something to open the door with' motivates the question: 'Then how do I know what I have to look for?' (153a6)	Feeling of confusion: 'I am in a muddle' (153a7)
(f) 175-7 (cf. 169-70) 'Sure enough, I was guided here' (175a4, b7), namely, , when copying an arbitrary doodle (175a1-2)	Interpreted in the light of A1 as: I had an experience of being guided, i.e. 'I experience[d] the because' (177a1)	When I say 'I was guided', I look for a characteristic 'experience of being influenced' (176a1)
(g) 175-7 Finding that I don't want to describe thus any experience I recall having (176a2, 177a2): 'But if I say what happened, I no longer find it characteristic' (175a6)	A1: Due to the tacit assumption that there <i>must</i> have been a characteristic experience, 'the idea of an ethereal, intangible influence arises' (175b8)	Feeling of dissatisfaction with <i>any</i> description of the event: 'It's as if I couldn't believe that I merely looked ... and drew a line' (175b3-4), i.e. intellectual unease
(h) 191-7 We understand the word, and its meaning lies in its use. (197a6)	When I understand a word, I grasp something that determines the future use logically, not merely causally (195a1) B: 'When one thing logically determines another, it somehow already is the other' (FF 355) tacitly presupposed in leap to conclusion: 'The future development must in some way already be present in the act of grasping' (197a5)	Sense of wonder articulated by: 'It is as if we could grasp the whole use of the word in a flash' (191/7a1), i.e. amazement – 'It's astonishing' – alternating with unease – 'It's queer' (cf. 197a3)

rather than in a characteristic feeling (177). In cases (c) and (g), by contrast, he jumps from correct findings to puzzling conclusions he throughout realises to be distinct from them. Finally, in the most complex cases (a) and (h), initial misinterpretations are exacerbated by an inference proceeding from them. In these last four cases, as in (f), Wittgenstein (ultimately) arrives at ideas he is throughout 'only' tempted to countenance, while he momentarily endorses the statement or question obtained in (b) and the opening move in case (a).

Most moves are in accord with the schema (A1). This autonomous cognitive habit thus manifests itself in two ways of distorting platitudes, which may mutually reinforce each other. *Against* this drive to misunderstand, and others like it, Wittgenstein wants to win through to such common-sense notions as that we usually understand a word we hear or say, that a sincere speaker will say 'Now I understand' because he understood, or that when copying a figure one is being guided by the original. He wants to get himself to apprehend these platitudes as what they are, without distorting them through inadvertent misinterpretation or mindless inference, in line with ideas he reflectively rejects.

But why should Wittgenstein want to pursue this, at first sight ridiculously modest, aim? For an answer, we need to turn from the distorted interpretations and conclusions to the 'behavioural and emotional consequences' both have regardless of whether or not they are momentarily being endorsed. The 'behavioural consequences' consist in the urge to hunt for snarks: for mental events of understanding and a characteristic experience of being guided. Partial yielding to these urges leads to disappointing findings – (b) and (h), analogous (c) – from which he leaps to puzzling conclusions. The behavioural thus exacerbate the crucial 'emotional consequences': the feelings of confusion, intellectual unease, and amazement that are unreasonable in the light of their pedestrian objects but intelligible in the light of the distorted constructions put on these. Wittgenstein analysed many of these distortions as such, himself (154, 169–70, 175–7). And he clearly implied that he found the problems he put down to such 'misinterpretation' deeply disturbing (111a1–2). He thus thought the return to the notions of common sense worthwhile because it would put an end to his unwarranted, but all too real, feelings of perplexity. As suggested above, these are the 'bumps' that make him appreciate the value of nonsense uncovered (119).

The comparison of his general predicament to that of Beck's patient may drive home this value. In the grip of the pertinent cognitive habit, the patient time and again put a wrong construction on others' remarks and demeanour, and on her own experiences. In automatic thinking, a compliment made with all signs of sincerity is interpreted as phoney, the yawn of a tired friend is regarded as a sign of boredom with her, her excitement before a talk is taken to reveal intuitive insight into her own ineptitude ('I can tell I'm going to muck it up again'). Accordingly, she regularly felt wretched, rejected, or afraid. Her autonomous cognitive habit thus turned her, in fact, reasonably

successful life into a series of apparent failures, and made her feel stuck in an endless mire of misery. While arguably due to different causes and manifestly accompanied by a far higher level of insight into it, Wittgenstein's predicament is structurally similar, and almost as disturbing: time and again in philosophical reflection, his autonomous habits of thought turn facts and findings that are actually rather pedestrian into a series of apparent puzzles, and make him feel bogged down in a vast mire of mystery. To escape from this disturbance constantly refuelled by drives to misunderstand is the declared ultimate aim of his philosophical work: 'Thoughts at peace [literally: Peace in the thoughts]. That is the goal someone who philosophises longs for' (MS 127, 41v: 4.3.1944; CV 50).

7 Wittgenstein's therapeutic approach

This aim can be pursued with more or less ambition. The greater ambition would be to prevent the recurrence of the thoughts (under 'B') that disturb him by engendering feelings of confusion, amazement and unease (under 'C') in the face of the familiar (under 'A'): to break the habit of inadvertent misinterpretation and mindless inference, in line with assumptions one rejects. The more modest aim would be to shed the feelings of perplexity the resulting thoughts hitherto went with: to learn to put on these thoughts a common-sense interpretation on which they no longer seem puzzling. In a severe case like Wittgenstein's, a realistic aim presumably lies half-way between the two: to expose and weaken the autonomous cognitive habits responsible for the various eruptions of perplexity, while learning to live without puzzlement with what remains of these habits in spite of all efforts. This is the bipartite task we shall see Wittgenstein address in sections 138–97.

At the 'local' level, he sometimes tries to place a common-sense interpretation on puzzling autonomous thoughts he did leap to. This move is at its most involved in the case of the crucial thought: "But I don't mean that what I do now (in grasping) determines the future use *causally* and as a matter of experience, but that in a *queer* way, the use itself is in some sense present." – But of course it is, "in *some* sense" . . . the sentence only appears queer when one imagines a different language-game for it from the one in which we actually use it' (195), namely, when one wants to use it as a description of some sort of process (196). It is idiomatic German to say that a use was 'present to one' when hearing or uttering a word. (That only the German use of the word was thus present – '*Mir war nur der deutsche Gebrauch von "originell" gegenwärtig*' – may be my excuse for calling a dress 'original'; for this reason, it did not occur to me that the English hostess might regard as an insult what was meant as a compliment.) In this sense, it is of course a manner of use that is present to one. But if we want to give the sentence 'The future use was present to me in the act of grasping' a reading that is less than puzzling, we can still render it as, 'The manner in which I would generally

use the word from then on was present to me when I first came to understand it'. We can, thus, dispel the unease induced by a puzzling autonomous thought that occurs to us, by coming up with an interpretation of its expression, which is in line with common usage. Presumably, this is what Wittgenstein does when pointing out that the expression of a thought perplexing him could be interpreted in such a way, as he repeatedly does in a more straightforward manner (here: 191a3, 197a2–3). This takes him from 'apparently queer' thoughts to notions of common sense and the conclusion, 'But there is nothing astonishing, nothing queer, about what happens' (197a4).

At a higher level, he employs precisely the techniques cognitive therapists use to identify and 'modify', i.e. weaken or break, relevant autonomous habits of thought. When they are not sustained by a powerful psychological motivation, such habits can be significantly weakened 'simply' by eliciting and refuting, again and again, autonomous thoughts and tacit assumptions manifesting them – as the success of Beck's cognitive therapy has amply proven. The two main moves are typically supported by giving two kinds of accounts: alternative accounts of the topics of the distorted thoughts at issue, and accounts of their formation that render intelligible how the subject 'could think such things' without implying the truth of these thoughts. All these moves are as straightforward as the 'trivial means' Wittgenstein took to free us from 'profound philosophical disquiets' (VW 70). Indeed, in sections 138–97 he makes these very moves time and again, while seeking out precisely the topics his thinking on which is most affected by the present habits. We will now consider and explain first the principal moves of exposure and refutation, and then the two moves made to support them. Wittgenstein has to deploy all these moves, in order to weaken his (in our sections) most persistent habit: the habit of applying the schema (A1). His other habits require a less full treatment. I therefore focus on his struggle with the 'toughest' case, in sections 147–80. Immediately before (139c–46) and after (181–3), he addresses (A2). Then, in section 184, he turns to (B).

1 *Exposure* proceeds through logical reconstruction of one's reflections, and probing for autonomous thoughts. Tacit assumptions can be exposed through logical reconstruction of pertinent reflections, which reveals leaps of thought and identifies unstated assumptions that would render them valid. Thus, Wittgenstein uncovers, in 154a, his previous leap of thought, from 152b2–4 to 153a1. Autonomous thoughts can be elicited by attending to what one feels inclined or tempted to say, at various points that are illustrated by the first three autonomous thoughts on the list above (in section 4). Wittgenstein attends to what he is tempted to say with approval when making a pertinent leap of thought (140a1–2 on 139d4), in protest when making a statement at odds with his habit's theme (such as the protest of 147a1–2 against 146b6), or in answer to a question (145b3=146a3) a correct answer to which (145b4) he feels inclined to interpret tendentiously (146b5–6).

Wittgenstein engages more in probing for autonomous thoughts than in logical reconstruction. Such probing requires some subtlety. Naturally, a subject in the grip of an autonomous habit, of a schema and a model to fill it in, will feel the inclination or temptation to say things that instantiate the relevant schema most strongly when dealing with topics to which the corresponding model can be applied in the most straightforward manner. In other words, autonomous thoughts manifesting the habit will be more salient on these topics than on others. Thus, consider the model (A) of a mental event: the workings of an inner mechanism. This model applies most straightforwardly if there is something we can readily conceive of as an 'input' that sets the mechanism going, and an 'output' generated by its workings. This we have in the cases of reading, taking down a dictation, copying, all regardless of understanding. In brief, in the cases which Wittgenstein, *prima facie* surprisingly, lumps together under the label of 'reading' (156a2), the 'input' consists in the written signs to be read out, the sounds uttered by the dictating teacher, the figure to be copied. The words uttered in reading, the signs written down upon dictation, the figure copied, then, are the 'output' (cf. 170c2). Contrast this with understanding. When we understand a sentence we hear, it may be tempting to think of the words uttered as the 'input'. But what is the 'output'? And if it is something we say ourselves, the sounds we produce might be thought the 'output' – but for what 'input'? Here, no answers suggest themselves naturally (but have to be cooked up). It is, therefore, no surprise that Wittgenstein reports no autonomous thought on understanding that clearly instantiates (A1), untainted by (A2), even though he frequently makes leaps of thought that tacitly rely on it; and that he reports a host of such thoughts on 'reading'. This lends his apparent digressions their point and purpose: to expose relevant autonomous thoughts, he repeatedly moves to the topics where they are most salient. When struggling with habit (A1), he moves to 'reading' (156–78); when dealing with habit (B) he moves to the topic of machine symbolism (193–4), where the model (B) of the ethereal machine applies in the most straightforward manner.

In the case of his thoughts on reading, Wittgenstein has to overcome a characteristic obstacle before he can elicit thoughts capable of refutation. He has to check a temptation to move from his first idea that the difference between true and merely apparent reading is made by a conscious mental act (156e) to an idea incapable of straightforward refutation: the idea that the difference could be made by reference to an unconscious cerebral process, accompanied by reluctance to take any other way of making the difference seriously (156g5). To check that temptation, Wittgenstein simply clarifies what it is that tempts him, and what he would be doing in yielding to the temptation. Having provided a context in which he feels tempted to voice the idea (158a1–3), he brings out that what tempts him is the a priori claim that a cerebral mechanism *must* make the difference between true and merely apparent reading – made in complete ignorance of any relevant empirical fact

(158a4–7). But, he concludes, if it is a priori, it merely articulates a 'form of representation' we find 'compelling' (158a8, my translation). He would merely rule out one manner of speaking and thinking in favour of another (cf. Z 442), on no other grounds than that he found the latter compelling, namely, found it compelling upon having to 'admit' that his first idea, involving a conscious mental activity, is false (156g1–2). This reflection is enough to make Wittgenstein return to thoughts revolving only around this first idea, which he proceeds to elicit.

2 *Refutation*: Wittgenstein then refutes these thoughts in different ways, depending upon their content. The bulk of the discussion of reading focuses on a sequence of autonomous thoughts in which each cedes only as much ground as made absolutely necessary by the refutation of the previous thought:

- (1) 'The one real criterion for anybody's reading is the conscious act of reading, the act of reading the sounds off from the letters' (159a1).
- (2) [Even if we do not seize on this conscious act or process as a criterion to distinguish genuine from merely apparent reading, still] 'reading is a quite particular process': We merely need to 'read a page of print' to 'see that something special and highly characteristic is going on' whenever we read (165a1–2).
- (3) [Even if no special or characteristic trait I am conscious of is shared by all cases of reading, still] 'I feel [in various different ways perhaps] a kind of influence of the letters working on me, when I read' (169c3).

The first, Wittgenstein considers a 'grammatical' claim. (2) and (3) formulate psychological claims. Accordingly, he addresses the former by considering how competent speakers would describe conceivable situations, and refutes the latter by empirical 'experiments' (169c7) with himself in the dual role of subject and observer. In both cases, it is worthwhile to bring out the subtleties of his proceeding that has been cast all too often into coarse moulds that will not fit.

To fully convince himself of the falsity of (1), Wittgenstein first clarifies its content by reminding himself of the various sensations characteristically involved in reading (159). Then (160a), he conceives of a case in which we would describe the subject as 'reading' even though he has none of these sensations but rather others, characteristic of reciting something from memory. (Upshot: we do not insist upon the former sensations as a necessary condition of reading.) Next (160b), he conceives of a subject who, under the influence of some drug, pronounces various arbitrary squiggles as if they were signs of a familiar alphabet, and does so with all the outward signs and

sensations characteristic of reading. Some competent speakers, Wittgenstein notes, would say that he reads, others that he does not. (I.e. the sensations adverted to are at least not universally acknowledged as a sufficient condition of reading.) More importantly, he finally goes on to develop the case in a way that makes clear that the outward signs, rather than the sensations, incline those speakers who are so inclined, to describe the subject as reading: if he goes on to consistently pronounce each squiggle in the same way, presumably most competent speakers would say that he makes up an alphabet on the spot and reads accordingly. This makes us aware of an implicit linguistic decision at odds with (1): the moment we consider concrete cases, we are manifestly inclined to use systematic responses, rather than characteristic sensations, as a criterion of reading.

The next thought (2) has two parts: there is something characteristic about the experience of reading (165–6); and whenever we read, one particular process takes place that we recognise (due to that characteristic experience, one will first think) (167–8). In response to the first idea, Wittgenstein teases out what he is inclined to regard as characteristic about the experience of reading: "The words I utter *come* in a special way." . . . They come of themselves . . . the spoken words as it were slip in as one reads' (165). To refute the idea thus rendered more tangible, he then makes an experiment: he first reads a letter and then thinks up a sound while looking at a flourish, to determine whether some such experience distinguishes the genuine first from the merely apparent second case of reading. He finds that he cannot grow aware of any experiential difference in the way the words came (166a3–12). To counter the second part of the idea, 'when we read one particular process takes place, which we recognise' (167a2), Wittgenstein reminds himself of some qualitative features of experiences had in reading, and states that none of them is in evidence in all cases of reading (168). This amounts to a refutation of the idea considered, if the qualitative features enumerated are all those he 'recognises', and if the final statement is supported by actual self-observation in various cases of reading – as his approach to (3) suggests. To refute the thought that 'I feel a kind of influence of the letters working on me, when I read' (169c3), Wittgenstein proceeds in two steps. First, he pins down the feelings or experiences that can, indeed, be aptly described as 'experiences of influence etc.' (171a), and then reads a few sentences in print as he usually does when he is not thinking about the notion of reading (171b). He seems to observe that in doing so he does not have any of the experiences previously identified, and exhorts himself not to try to get around this finding, e.g. by saying that he had them 'unconsciously' (ibid.).

3 *Accounts of thought-formation*: To get himself to fully accept these refutations, i.e. to shed all temptation of getting around them by hunting for such snarks as unconscious experiences of being influenced or guided, Wittgenstein offers accounts of how he could think such things, that expose

his autonomous thoughts as 'fictions' (166a2) or 'imagination' (170c1) by rendering them intelligible as misguided interpretations of various kinds of things. In the discussion of (2), he clearly implies that he interpreted a difference in the situation of utterance as a difference in the way the words came (166a13–17), and the uniformity in the appearance of pages of print (the case of reading adverted to when first venting the pertinent idea, in 165a1–2) as the recurrence of a characteristic trait in the experience of reading (167). He then explains how (3) arises from the experiential difference he notices when, and only when (171b), he makes the experiment of first reading some letters of print and then looking along a line of flourishes written down for this purpose (169a2). The difference is this: 'when I see the letter it is automatic for me to hear the sound . . . inwardly . . . and I pronounce the letter more effortlessly when I read it than when I am looking at "§"' (169c6). In (3), he interprets this experiential difference as one between an influence, and the lack of one (170a2–3). Wittgenstein pins down the situation in which this interpretation is particularly appealing (170b): when reading slowly, with a particular philosophical question in mind (What happens when we read?), which is about an event or process. We may add that the thought thus betrays the influence of both schema and model (A). In line with (A1), Wittgenstein's thought generalises unduly and has it that whenever he reads, he has the feeling he is tempted to regard as one of influence. And it foists the mechanistic model of mental events on the supposedly ubiquitous feeling, as Wittgenstein finally spells out: 'when I speak of the experiences of being influenced . . . that is really meant to imply that I as it were feel the movement of the lever which connects seeing the letters with speaking' (170c2).

Here, and a moment later (in 175–7), Wittgenstein applies to his autonomous thoughts about *reading* an idea he had, in the *Frühfassung*, explicitly explained with respect to *understanding* (to which topic, however, he does not apply it, presumably for lack of pertinent autonomous thoughts):

Consider the pronunciation of a word as its spelling presents it. How easy it is to persuade oneself that two words – e.g. 'fore' and 'four' – sound different in everyday use – because one pronounces them slightly differently when one has the difference in spelling directly in view. Comparable with this is the opinion that a violin player with a fine sense of pitch always strikes F somewhat higher than E sharp. Reflect on such cases. – *That* is how it can come about that the means of representation produces something *imaginary*. So let us not think we *must* find a specific mental process, because the verb 'to understand' is there and because one says: Understanding is an activity of mind.

(BFF 172, between later sections 154 and 155;
kept as Z 446b)

This subtle idea is not to be confused with the more familiar notion that particular linguistic expressions in common use are philosophically misleading to all and sundry, or at least to anyone engaging in abstract reflection. In Wittgenstein's present example, the subject actively 'persuades himself' of his erroneous idea, and is not passively misled by the different spelling. Second, in doing so he seizes on something he actively does in quite specific situations which involve no explicit abstract reasoning: pronouncing two words when having in mind the difference in spelling, or reading slowly when asking oneself what is happening when one reads. Third, what Wittgenstein regards as the crucial 'means of representation' are not only symbols but philosophical statements (instantiating our schema A): 'Understanding is an activity of mind' or 'True reading presupposes the setting up of a connection in the reader's mind, or brain' (the 'form of representation' he finds tempting in section 158).

What the self-analyses in line with these ideas actually reveal, thus, are misinterpretations of things he says ('I'm being guided' (175–7)) or of experiences he enjoys under quite specific circumstances (169–71), that betray an active *drive* to misunderstand: the autonomous cognitive habit pinned down by schema and model (A). The reference to 'means of representation' obviously yields no explanation of why Wittgenstein has this habit. To take up his present example: when their attention is drawn to the different spelling of 'four' and 'fore', most language students merely curse the lack of systematic correlations between spelling and pronunciation – and make no effort to persuade themselves of anything. Only those (all too) ready to be misled will be unduly impressed by this or that feature of the pertinent 'means of representation'. So what function do Wittgenstein's self-analyses actually fulfil, if they do not explain why he came to have the thoughts at issue? Arguably, the same function that the overtly non-explanatory 'accounts of thought-formation' offered in the course of cognitive therapy serves: they 'merely' reveal the extent to which the subject leaps to wrong interpretations of innocuous 'activating events', spontaneously, unthinkingly, and without warrant. This insight into the nature of his own thinking is to nurture in the subject a mistrust of his interpretative urges, and to help him accept the refutation of the thoughts that result from them, here: to counter the urge to conclude that, all the same, there *must* be a specific mental process of understanding (even though it is strangely elusive), that there must be something characteristic about the experience of reading or copying (which he cannot pin down), etc.

4 *Outline of alternatives*: Not only the final but also the opening move of the discussion of reading makes good sense, the moment it is regarded as fulfilling the function of a supporting move common in cognitive therapy. In section 157, Wittgenstein points out how the difference between genuine and merely apparent reading *can* be drawn without reference to any 'inner' event, namely,

by reference to further responses to written signs, in other situations. He provides a context in which this proposal strikes us as attractive. If we trained people to serve as 'reading machines' we would, indeed, care first and foremost about whether they consistently respond correctly to given signs. Finally, Wittgenstein derives from either alternative a consequence that can be tested against our ordinary way of speaking. If we used a certain experience or inner mechanism as a criterion of reading, we could say of a particular word that it was the first word the subject read; when we refer exclusively to his overt performance, we cannot say this. The tacit upshot is: as we indeed do not say this, we must commonly distinguish genuine from merely apparent reading without reference to any 'inner' event but, rather, in something like the way he imagined for the 'reading machines'. This clearly will not do as an analysis of our ordinary concept of reading (which 156a2 anyway tells us will not be discussed). But it does fulfil a function 'alternative accounts of the topic at issue' serve in cognitive therapy: by revealing that things do not *have to be* as the subject thought, it drives home that the subject's spontaneous thoughts on this topic (reported in 156e and g) already constitute an interpretation in the light of unquestioned background assumptions. Appreciation of this point is to help the subject expose and soberly assess pertinent autonomous thoughts. To which bipartite exercise Wittgenstein then proceeds.

He thus makes all the four major moves of cognitive therapy: exposure and refutation of pertinent autonomous thoughts, supported by accounts of their formation and outlines of alternatives to them. In the present context, all moves manifestly fulfil (some of) the functions cognitive therapists commonly deploy them for. While his other cognitive habits require less full treatment, Wittgenstein's struggle with his most persistent habit of thought amounts to a full-blown cognitive (self-)therapy.

These findings allow us to make good sense of the apparent marks of ineptitude in the text, noted at the outset. So, far from haphazardly jumping from topic to topic, Wittgenstein addresses precisely those topics his thinking on which is affected by the habits or 'drives' he wants to expose and weaken. He arranges them in quite neat order: sections 139–83 treat the habit of thinking in line with the mentalistic picture (A), sections 184–97 deal with the actualistic picture (B). The former treatment divides neatly into two: the struggle with (A1) in sections 147–80 is embedded in the treatment of (A2), in 139–46 and 181–3. The reprise makes sense: sections 181–3 deal with (A2) in the new guise in which it re-emerges after the treatment of (A1). Second, both of his apparently wild digressions are well-motivated: the models of the two pictures at issue apply most directly to reading and the use of machine-symbolism, making these topics particularly suited for eliciting pertinent autonomous thoughts. Third, the feelings of perplexity in the light of sudden understanding, surfacing in sections 138–9, 151–3 and 191–7, are due to the interplay between different autonomous habits of thought, and

will persist as long as these habits have not been weakened. Which can only be done gradually, by eliciting, refuting and analysing a *host* of pertinent, related but diverse, autonomous thoughts. Hence, it is no sign of ineptitude that thoughts about sudden understanding, supposed and genuine, are being discussed again and again.

8 Wittgenstein's philosophical relevance

This exegetical success recommends the present account of Wittgenstein's predicament, aim, and approach. He is in the grip of autonomous habits of thought that have him jump, in line with assumptions he explicitly rejects, to ideas that are not even faintly sensible but induce in him a sense of wonder or confusion, whether or not he momentarily endorses those ideas. Either way, these feelings are pathological in the sense that they are utterly unreasonable in the light of his own stable, reflective beliefs. His aim is to shed these pathological feelings of perplexity, constantly refuelled by 'drives to misunderstand': to achieve 'peace in the thoughts'. To do so, he exposes and weakens the pertinent 'drives', i.e. habits, employing what later became the core techniques of cognitive therapy. To take the mystery out of what distorted thoughts keep recurring in spite of these efforts, he puts on them interpretations in line with our ordinary ways of talking.

If successful, the result of Wittgenstein's auto-therapeutic endeavours will be that *he* is free from bogus puzzlement. But no insight (except into the nature of his own personal problems) will have been gained, no truth about the world established. This lends new content and force to the question of relevance that has puzzled so many readers of Wittgenstein: why should his cognitive self-therapy, even if successful, be of interest to anyone else – and, in particular, to any other philosopher?

In a nutshell, the answer is this: the root of Wittgenstein's predicament is by no means unique. Most serious philosophers are driven by autonomous habits of thought. Systematic misinterpretations and leaps of thought, in both automatic and controlled thinking, generate bogus mystery, articulated by pseudo-problems philosophers then waste their lives working on. Many philosophers are in need of therapy to save them from unwarranted perplexity and waste of effort. Wittgenstein's text is therefore instructive to them, to the extent to which it helps them gain insight into their own predicament and furnishes them with a pertinent therapeutic approach. I develop these ideas in detail, in my forthcoming *Philosophical Delusion and its Therapy* (Fischer 2005). Here, a brief illustration may at least render them intelligible.

Consider the metaphysical vision of phenomenalism. It rests squarely on two exemplary philosophical intuitions: that all we perceive are sense-data, and that we might as well 'cancel through' anything imperceptible 'behind the representationalist's veil'. To support the all-important first intuition, its proponents typically adduce the 'argument from illusion' that starts out like this:

Consider . . . [a straight] stick which is refracted in water . . . [I]t must be assumed that the stick does not really change its shape when it is placed in water . . . Then it follows that at least one of the visual appearances of the stick is delusive; for it cannot both be crooked and straight. Nevertheless, even in the case where what we see is not a real quality of a material thing . . . we are still seeing something; and . . . it is convenient to give this a name. And it is for this purpose that philosophers have recourse to the term 'sense-datum'. By using it they are able to give what seems to them a satisfactory answer to the question: What is the object of which we are directly aware, in perception, if it is not part of any material thing?

(Ayer 1940: 4)

This question clearly presupposes what is to be argued: that we are not seeing 'part of any material thing'; and it is particularly remarkable in view of the fact that it has been assumed at the outset that we do see an exemplary material thing, namely, a stick – and, presumably, both the part above and the part in the water. Logically valid argument leads up to a fancy and unnecessarily cautious formulation of the platitude that when immersed in water the stick appears crooked, even though it really is straight. From this, a mere leap of thought leads straight to the conclusion: that when something appears to us different than it really is, we are seeing something immaterial, a sense-datum. So far from supporting it, the 'argument from illusion' merely articulates a philosophical intuition, after pointing out a situation in which it strikes those who have it as particularly powerful. The same is true of many of precisely the most influential pieces of philosophical prose: brief but compelling stories misleadingly called 'arguments', such as, say, Descartes' 'argument from dreaming'. Philosophers' visions, metaphysical and other, tend to rest on intuitions unsupported by argument.

In the situations outlined by the stories passed off as arguments, these intuitions occur as automatic thoughts. When letting our thoughts drift in contemplation of the sight, with neither a purpose in mind nor any particular interest in the stick (or the water it is immersed in), we spontaneously want to say, 'Really, all I see is that crooked speck, which I cannot touch'. This automatic thought is then treated as a philosophical intuition to be honoured at all cost. From the assumptions that such 'delusive' are qualitatively indistinguishable from 'veridical' perceptions, and that perceptions that are qualitatively alike must have the same kind of object, it is inferred that all we ever see are specks, sense-data, rather than material things. In spite of the quite patent absurdity of these last assumptions, not even their remarkable consequence causes the proponents of the 'argument' to take a deep breath and question their sudden thought that they 'really' see only the crooked speck, not the straight stick. Thus, many philosophers are no strangers to automatic thoughts involving terrible distortions. The most salient of these are the

'philosophical intuitions' they regard as encapsulating deep, if partial, insights into profound general truths, 'truths' they duly proceed to derive from them, if necessary with the help of the most dubious further assumptions. The result of this distorted thinking is the articulation of a philosophical vision that is just that: a groundless *vision*.

The automatic thoughts at the beginning of the garden-path are typically part and parcel of an autonomous cognitive habit that biases our thinking from the start and to the end. As we have seen, we inadvertently rely on the assumption that we do not perceive anything material already in the course of the very 'argument' meant to establish this. And we may stick to this assumption, an instance of the present habit's schema, with peculiar persistence. Upon a moment's reflection on the initial situation, we may want to add: 'Of course, it is perfectly good English to say that we then see the stick, a material thing. But I see the speck somehow more directly, and it is all that I directly see.' The moment we become clear on what we mean by 'direct perception', we find we want to apply this new expression only where the object of perception cannot appear otherwise than it really is, and may realise that we are using it to talk about the appearances of things: to describe how they appear to us in particular situations, rather than how they actually are. But, of course, what appears to us thus still is the stick, and adherence to the 'intuitive insight' that all we see is a bent sense-datum, it turns out, amounts to no more than the insistence upon a new way of talking about how that material thing appears to us. Fine philosophers like A.J. Ayer rehearsed this line of thought, stressed its conclusion – and proceeded to interpret, conversely, our ordinary talk of material things as a convenient way of describing our sense-data. This interpretation is in line with the core doctrine that all we really see are such sense-data, which they had just found unwarranted by what 'argument' they could adduce: yet another manifestation of the autonomous cognitive habit. The groundless vision is thus generated by an autonomous habit of thought that then keeps the philosopher under the vision's spell – even if he is in a position to know better.

Thus, the habit gives rise to unwarranted perplexity and may motivate pointless efforts. The perplexity is due to the mismatch between the vision inspired by the 'intuitions', and the more mundane ways of thinking and talking that the philosopher is unable and, perhaps, unwilling to shed. Thus, the phenomenalist wonders: 'How is it possible for us to use, every day, expressions that ostensibly refer to publicly accessible material things with causal properties, located in physical space?' (cf. Ayer 1940: 244) – a pseudo-problem only raised by his vision that all we actually see (indeed, all that actually exists) are transitory and private sense-data devoid of physical extension and causal efficacy. To remove what unease this clash between his vision and common sense may cause him, and at any rate to reconcile the two, he then rushes to develop a philosophical theory that purports to answer that question. Which is, of course, rather a waste of time and effort: as his vision

is entirely groundless, there is no point in reconciling it with common sense (or anything else, for that matter).

If it conflicts with our reflective beliefs, one should instead try to shed the vision, so as to put an end to both the unwarranted perplexity and the waste of time and effort. But, as phenomenism illustrates, a philosopher's vision typically is not under his control. Powerful autonomous habits of thought first force it on him, and then keep him under its spell. They force it on him through a series of automatic thoughts exacerbated by leaps in controlled reflection. And they may keep him under its spell even if he has seen through those supposed 'insights'. When affected by such an autonomous habit, a philosopher is in need of a therapy that weakens or breaks it. Not suffering from a compulsive disorder, he is of course free to decide to put an end to his pointless endeavours, any time; but he needs therapy to be able do so without regret and the feeling of a task left undone, of perplexing puzzles evaded. Arguably, quite a few serious philosophers were, and are, in this situation.

9 Wittgenstein's therapeutic turn

One of these philosophers, of course, is Wittgenstein. Most of the *Tractatus* derives its point from the metaphysical vision of a world's essence mirrored by logic, resting on such intuitions as that a proposition represents a state of affairs in the way a model used at court pictures an accident, intuitions to be honoured even at the price of inventing 'logically proper names' and postulating sempiternal simples of which one knows nothing. The most important, and most novel, move Wittgenstein later made is to have distanced himself from such automatic thoughts: to have conceptualised them not as philosophical intuitions to be honoured at all cost but as unfounded impulses to say things that are, all too often, not even faintly sensible, as one realises the moment one does not deliberately dismiss one's common sense, driven by an 'urge to misunderstand'. In these cases, one needs to expose the impulses as what they are, automatic thoughts that go unquestioned, and to weaken them to a point at which one no longer feels compelled to 'honour' them. This, I think, constitutes the most fundamental and revolutionary aspect of the later Wittgenstein's manifold reorientation of philosophical work: his 'therapeutic turn'. Wittgenstein, himself, gradually became clearer on its content and eventually articulated it by comparing 'the philosopher's' treatment of a question with the treatment of an illness (255): 'What we are "tempted to say" in such a case is of course not philosophy; but it is its raw material. Thus, e.g. what a mathematician is inclined to say about the objectivity and reality of mathematical facts, is not a philosophy of mathematics, but something that philosophy would have to *treat*' (254a4–5, my translation) – like an illness, namely, by weakening the inclination to say those things that give rise to pointless questions.

After this reorientation of his philosophical efforts, in the very early 1930s, Wittgenstein still had to struggle for considerable time against the cognitive

habits he had indulged when working on the *Tractatus*. Much perplexity continued to be engendered by the autonomous habits that persisted in his philosophical reflection, namely, by the mismatch between them and his ordinary ways of thinking and talking. At this point, he differed in three respects from most other philosophers. He no longer mistook the elaboration of a groundless vision for the pursuit of truth, combining insight with the decency – and strength – to acknowledge it. As a result, he abstained from constructing – pointless – philosophical 'theories'. But, even so, he felt the perplexity much more keenly than most philosophers. At root, however, his predicament is the same as theirs: he is in the grip of autonomous habits of thought that confront him with bogus problems. Any philosopher in this not uncommon situation can, therefore, profit from the *Investigations*' record of Wittgenstein's efforts to liberate himself from autonomous cognitive habits he had earlier indulged. From it we can derive insight into our own predicament, and a therapeutic approach to cope.

To conclude: all too often, philosophy is toil to conquer lands that have in them no gain but the name. Therapy is to ease the urges that drive a philosopher into these wastes: to make him resist them, without misplaced regret. In these parts of our subject, the alternative to the modesty of therapy is the valour of Don Quixote. We can do our duty towards ourselves, or cultivate a vision's gentle madness. Wittgenstein's work is instructive for those who share both the root of his predicament and his preference of duty over self-deceit.³

Notes

- 1 Reference to section (180), paragraph (a) within section, and sentence (1) within paragraph, of the German text. The parsing into sentences is not always obvious. But a rough indication seems better than an even rougher one.
- 2 Section 255 as written in 1945 (MS 116, 323) and typed into the final typescript (TS 227), typed up/compiled in late 1945/early 1946. No earlier did the relevance of 133d, jotted down in 1937/8 (MS 116, 186), become fully apparent to Wittgenstein. He appended it to that typescript on an extra slip of paper. Some further remarks of similar import date from 1944–7. The much noted earlier (1931) mention of psychoanalysis (FF 108, derived from BT 409–10), by contrast, is quite irrelevant in this context. It compares neither his problems nor his approach to anything treated or done in psychotherapy, but primarily elucidates a criterion of correctness for the attribution of certain lines of thought. (The most pertinent early related remark, VW 68–70 in the Waismann-papers, is probably of no evidential value either, as the editors take it to be 'mislocated' in the context in which it would imply an etiological claim.)
- 3 For helpful comments, I am indebted to Erich Anmereller, Cora Diamond, Peter Hacker and Stephen Mulhall, for the historical information in Fn. 2 to Joachim Schulte. My most important and long-standing debt is to Gordon Baker whose encouragement and criticism, through many years, helped me to develop concepts of philosophical therapy to a point at which they could profitably be employed in Wittgenstein-exegesis.

Bibliography

- Ayer, A.J. (1940) *The Foundations of Empirical Knowledge*, London: Macmillan.
Beck, A.T. (1976) *Cognitive Therapy and the Emotional Disorder*, New York: International Universities Press.
Fischer, E.J.D. (2005) *Philosophical Delusion and its Therapy*, London: Routledge.

PUZZLES ABOUT RULE-FOLLOWING

PI 185–242

Erich Ammereller

In this essay I shall offer some comments on Wittgenstein's remarks on rule-following in the *Philosophical Investigations* (185–242). I share the view voiced by many commentators that these remarks provide a key to the proper understanding of Wittgenstein's thoughts on the nature of language, meaning and understanding. However, I find myself in disagreement with the way this key role of Wittgenstein's remarks tends to be construed in a considerable part of the literature. The main reason for what often seems to me to be a misconstrual of the purposes of Wittgenstein's considerations is a tendency of many interpretations to neglect or disregard what, in my view, is at once the most striking and challenging aspect of the *Philosophical Investigations*, namely, the view they present of the nature of philosophical problems and their proper treatment. This tendency is by no means a peculiarity of the discussion of Wittgenstein's remarks on rule-following though, as I hope to show, it is more of a curiosity with regard to them given the frequency and explicitness with which he gives prominence to their methodological purpose. By focusing, more than has been usual, on this aspect of his investigation. I hope to throw a little more light both on Wittgenstein's rule-following considerations and on his later way of thinking.

1 Philosophy

Any reader of the *Philosophical Investigations* must be struck by what Wittgenstein says about 'the work of the philosopher' (PI 127):

It was true to say that our considerations could not be scientific ones. [. . .] And we may not advance any kind of theory. There must not be anything hypothetical in our considerations. We must do away with all explanation, and description alone must take its place.

(PI 109)