

that the senses of our expressions are to reach out to the world and to one another in a wholly objective, mind-independent, manner. On this view sense is to *determine* what is to count as the correct use of an expression, rather than being *constituted by it*. (Meaning is not use, but rather determines use.) And so understanding – knowledge of sense – must consist in an intellectual grasp of that which determines correct use. Hence understanding cannot consist in a bare practical capacity, or at least not in the way that the Socratic-fallacy-defence requires. On the contrary, a person's capacity to use an expression correctly is to be *explained in terms of* their understanding of it, the explanation being that they are capable of deriving a knowledge of the correct use of the expression from their knowledge of its sense.

It follows that logical objectivists cannot brush aside our challenge by claiming that they simply have the capacity to use their expressions correctly, and to recognise internal relationships between them, without knowing what it is that makes their usage correct, or that in virtue of which those internal relationships hold. For this would be to concede that they do not, after all, understand (know the senses of) the expressions of their language. If understanding must involve, for logical objectivists, some sort of conscious awareness of whatever it is which is supposed to determine correct use, then they cannot intelligibly claim that they nevertheless have no idea what the senses of their expressions might be.

Nor can this difficulty be avoided by denying that understanding need involve conscious knowledge of sense, claiming rather that our grasp of that which determines correct usage and the internal relations between expressions is unconscious. For this would undercut the essential normativeness of sense. Speakers are supposed to be able to appeal to the senses of their expressions in justifying the use which they make of them, and in correcting the usage of others. Yet they could hardly do this if their knowledge of sense were unconscious. Indeed, on such a hypothesis an unbridgeable gap would open up, between both actual usage and the sorts of justifications we offer of it on the one hand, and what would really determine correctness or incorrectness on the other. Since we would not be supposed to have any conscious awareness of sense, we should be incapable of knowing whether or not our actual normative practice was appropriately justified. Indeed, it would be possible to doubt whether we really attached any senses to the expressions of our language at all. But this sort of scepticism must be wholly unacceptable. For whatever room there may be for doubts about other people's minds and meanings, I cannot doubt but that I, at any rate, mean something by the words that I use.

15 Deconstruction: following a rule

In this chapter I argue against the logical objectivism of TLP, suggesting that this very critique may be found in the later Wittgenstein's discussion of rule-following.

(A) An argument in outline

The main argument against logical objectivism takes the form of a challenge to find something which can play the role of the sense of an expression, as logical objectivism conceives of it. In the next section we shall consider the various possible attempts to meet this challenge head-on. Here we shall consider the broad structure of the argument, supposing that the challenge cannot in fact be met.

Suppose then that logical objectivists can provide no answer to the question what it is to understand (to grasp the sense of) an expression. Indeed, suppose that they can find no object of awareness having the properties which they attribute to senses. What would follow from this? They might be tempted to claim that nothing follows, pointing out that it is precisely the fallacy sometimes labelled 'Socratic' to infer from the fact that speakers cannot say – cannot articulate – what their understanding consists in that they therefore possess no such understanding.¹ But in fact this line of defence is unpromising. For the standard way of explaining how the Socratic fallacy is indeed a fallacy is in terms of the distinction between knowing *how* and knowing *that*; the point being that speakers can know *how* to use a given expression correctly without knowing *that* the correct use of that expression is such-and-such. But this is not a distinction of which logical objectivists can avail themselves, in replying to our basic challenge.

The reason is that they cannot consistently equate understanding with a practical capacity: a matter of knowing *how*. For remember their idea is

The upshot is that if it could be shown that we are none of us aware of anything which can play the role of the sense of an expression, as logical objectivism conceives of it, then it would follow (supposing logical objectivism to be true) that we none of us understand or mean anything by the expressions that we use. For understanding must, on such a conception, be essentially a matter of consciousness of senses. This would reduce logical objectivism to absurdity. We should then be faced with a choice of either rejecting the notion of sense altogether, or finding some alternative conception of it.

(B) Candidates for sense

Some suggestions can be ruled out straight away. We can discount mental images as being hopelessly implausible to account for the meanings of most words in an idiolect, let alone the meanings of sentences. And we can rule out thoughts as begging the question, since on the account of thinking which I (and Wittgenstein) believe to be correct, it is precisely one of those activities which needs explaining, on the logical objectivist account. For if thinking consists in the significant employment of sign-tokens, then we shall want to know how these signs are to reach out to reality in a manner independent of the dispositions of the thinking subject.² Fregean thoughts (*Gedanken*) can also be rejected as candidates since they fail to figure in the phenomenology of understanding. When I judge that an object is red I am certainly not aware of grasping a mind-independent abstract entity, nor is my use of the component terms in any way guided by awareness of such a thing.³

As I have noted at various points, *TLP* appears to be committed to a form of modified conventionalism, the sense of an expression being none other than the rules and conventions which determine (objectively, not conventionally) its use.⁴ So we then need to ask how rules are present to consciousness. Some, of course, may take the form of explicit definitions. But this only pushes the enquiry further back, on to the sort of understanding we possess of the other terms involved. The most basic mode of explaining a rule must be by means of examples and training. But in fact it is clear that nothing we can give someone in the course of such training (or give ourselves by way of an exhibition of the rule) can determine by itself a unique pattern of application. In which case there is nothing here which can play the role of a logical objectivist sense.

For example, suppose that in explaining the rule for the use of the term 'red' to someone we exhibit a wide range of samples of shades of red, as

well as samples of things that are not red. It is clear that a trainee could (at whatever stage) go on from this in a manner which we would consider perverse. For there will always be some shades of blue (for example) which have not been exhibited in the course of the training, and the trainee may naturally be inclined to respond to our explanations by classifying these together with red. And even if we were to go systematically through every distinguishable shade in the colour-spectrum, still our samples would have occupied some determinate position in space and time, and it would be possible that the trainee should cotton on to a rule which is some function of these – for example a rule which we might express by saying: 'is red' applies to an object if it is red in 1990, blue in 1991, yellow in 1992, and so on.⁵ Since this must always remain a possibility, there can be nothing in the training which can reach out to determine a unique pattern of application in a manner independent of our dispositions.

Since nothing which can be exhibited in public use can play the role of a logical objectivist sense, we might be tempted to appeal to some object of private awareness, most plausibly an intention. It might be said that a speaker's understanding of a rule will consist in a certain intention, namely a commitment to employ the term in question in a particular manner. Indeed, this can seem quite a plausible line for logical objectivists to take.⁶ For we not only appear to have non-inferential (perhaps even incorrigible) awareness of our own intentions, but we also seem incapable of error in judging whether or not an action accords with our intentions. For if I do not misapprehend what I am doing, and sincerely believe that I am doing what I am intending to do, then so I am. These facts might then be taken to show that intentions are objects of direct awareness which reach out to the world to determine what actions would or would not accord with them in such a way as to leave no room for misinterpretation – which is just what logical objectivism requires.

What gives the lie to this use for our ordinary notion of intention, however, is the phenomenology (or rather lack of it) of our awareness of our intentions. For to be conscious of an intention is often just a matter of finding oneself thinking about the intended action with one's mind made up. Perhaps I ask myself 'What was it that I intended to do today?' I may then recall having made up my mind to weed the garden, or I may simply think 'I shall weed the garden' in the knowledge that this is what I had made up my mind to do. Moreover, our awareness of an intention is quite unlike consciousness of a pain. For the former is wholly lacking in phenomenological content: it does not *feel like* anything to have the intention of weeding the garden.

There are of course cases where I have an intention, and am conscious of having it, without it being formulated in thought at the time. For example, while engaged in an argument I might know that I intend to hit my opponent if he says another word, without any thought of the form 'If he speaks again I shall hit him' going through my mind. But this in no way lends support to logical objectivism. For consider what occurs when I later recall such an intention. In one sense all I remember are the details of the whole situation – what was said, what I was thinking, and so on. Remembering my intention does not involve recall of a conscious state which existed alongside of these, but is rather a matter of having non-inferential knowledge, as I look back, that if another word had been said I would have turned to violence. I certainly do not have access to any conscious state which is capable of reaching out to the world in a manner independent of my capacities and dispositions, which is what the logical objectivist requires.

In fact the most plausible account of the features of intention to which the logical objectivist appeals is provided by a functionalist theory of the mind. An intention is a state apt to cause the intended action, conscious intentions also being apt to emerge in conscious thinkings with the same content, either at the time or later in memory. This is why knowledge of our intentions is non-inferential. It is also why we cannot in general be mistaken about whether or not an action accords with our intentions. For the intention which causes the action will also, if it is conscious, be apt to emerge in a thought which describes just such an act; so if I do not misapprehend what I am doing there is no further room for error. If this account is correct, then it follows immediately that intentions are not fit to serve as logical objectivist senses. For since they presuppose a notion of thought, just the same problems would arise about them. But I do not here need to rely upon the correctness of the account. It is sufficient to have established that knowledge of our own intentions is not a matter of awareness of a state which can reach out to the world in a mind-independent way.

We are then unaware of anything which can play the role of a logical objectivist sense.⁷ But neither can it help to maintain that the notion of sense is *sui generis*, insisting that nothing further can be said about what I am aware of when I know the sense of an expression except that I am aware of its sense. For this is to forget that the challenge to the logical objectivist applies particularly in the first person singular. We are each of us invited to introspect: when I mean something by a word or sentence, is there anything in consciousness which can determine a unique pattern of

use in a wholly objective mind-independent way? And the answer to this is clearly negative. In which case, since our challenge cannot be met, logical objectivism must be false. It turns out to be nothing more than a kind of mythology of symbolism. It is a picture of objectivity which, despite its attractiveness, vanishes from our grasp as soon as we try to render it concrete and substantial.

(C) The argument and *PI*

Readers familiar with *PI* will have noted in the above argument many echoes of Wittgenstein's famous discussion of rule-following. This is no accident, since the argument was constructed *via* reflection on that discussion, and since I take Wittgenstein's critical target to be none other than the logical objectivist conception of sense and relations between senses. Certainly that conception seems extremely close to the intended point of the metaphors of 'all the steps in the application of a rule having been taken in advance' (*PI* 188), 'a rule sending out infinitely long rails' (*PI* 218) and 'all the possible movements of a machine [as symbol] being already there in it in some mysterious way' (*PI* 193–4) which figure so prominently in the discussion. What Wittgenstein is wanting to criticise is the idea that in understanding one grasps something which is then able to reach out to the world in a mind-independent way, replacing this with the claim that understanding is a practical capacity: a matter of mastery of a technique (*PI* 182, 199).⁸

Not only is Wittgenstein concerned to attack logical objectivism, but the moves he makes closely parallel those of the argument above. Thus he criticises images as candidates for the meanings of words (*PI* 139–41), as well as arguing that no amount of training in the use of examples can determine independently of our reaction to it a unique pattern of application (*PI* 185–7). He also invites us to introspect, to consider whether there is anything in our consciousness which is able to reach out beyond the examples used in training (*PI* 208–9). He even tackles head-on the suggestion that a speaker's grasp of a rule might reside in their intention. This is covered rather briefly within the parts of *PI* most explicitly concerned with rule-following (*PI* 197 and 205), but he returns to the topic later, exploring it with great subtlety (*PI* 630–93). One clear theme of this later discussion is that intentions are not introspectable items in consciousness with a distinctive phenomenology; there is also a suggestion that conscious intentions are states which are apt to emerge in thinkings with the same content, either at the time or later in memory.⁹ But in any case he clearly rejects any

conception of intentions which might fit them to play the role of logical objectivist senses.

Another way of supporting my reading of the *PI* discussion of rule-following is to indicate how it is at least consistent with those interpretations which have recently attracted most attention. For example, note that it is closely related to the interpretation provided by Wright, who sees Wittgenstein's attack as being focussed on the objectivity of conceptual commitments.¹⁰ But as will emerge shortly, I differ from Wright in thinking that it is far from clear that Wittgenstein must then be committed to rejecting the objectivity (let alone the investigation-independence) of truth. My interpretation also agrees in detail with the one provided by McGinn, except that he makes no attempt to motivate the arguments he attributes to Wittgenstein, never noticing that we can understand their point and significance if we see them as directed against logical objectivism.¹¹ My reading even agrees substantially with Kripke's, except that he sees the whole notion of meaning as being under attack. In fact Kripke himself, in expounding Wittgenstein, assumes that meaning must have indefinite significance as well as being available to consciousness – which are the distinctive marks of a commitment to logical objectivism, as we have seen. What he fails to notice is that by equating understanding with a practical capacity we can retain the normativity of meaning without being committed to the objectivity of semantic relations.¹²

If these remarks are sufficient to establish that logical objectivism is the intended target of the discussion which is generally recognised to constitute the very heart of *PI*, then we have finally found our way to a basic locus of conflict between Wittgenstein's early and late philosophies. For as we have seen, logical objectivism is crucially involved in many of the most controversial aspects of *TLP* – in its conception of metaphysics, in the idea of a programme of analysis, in the commitment to eradicate vagueness, in the argument to *Simples*, and in the requirement that elementary propositions be logically independent of one another. If the attack which is mounted on logical objectivism in the following-a-rule sections of *PI* is successful, as I have argued it is, then all of these strands in his earlier thinking must collapse.

(D) What price objectivity?

If logical objectivism is to be rejected, then what becomes of the arguments of Chapter 4 which were left standing in its support, premised upon the objectivity of contingent and of necessary truth? Must belief in these kinds

of objectivity now be given up? If so, then our attack upon logical objectivism will have done a great deal more than undercut some of the more controversial doctrines of *TLP*, since a large part of common-sense belief will also have been undermined.

Consider first the objectivity of contingent truth. If we conceive understanding to be a practical capacity, rather than a grasp of something (a sense) which is able to reach out to determine a truth-value in a manner independent of us and our dispositions, then can there be determinate truths which outreach our capacity for verification? In fact it is by no means clear that a negative answer is forced on us. For the capacities held to be constitutive of understanding, which serve to project a sentence on to the world, are primarily those corresponding to its component words and their mode of combination. And these capacities may be such as to yield a determinate truth-value even in cases where we are incapable of verifying the sentence as a whole. In the same way, a subject may manifest their grasp of an objective truth-condition by manifesting the capacities which constitute understanding of the component words and their combination.¹³

For consider: according to anti-logical-objectivism, an understanding of a term is a multi-faceted capacity: namely to put that term together with others in such a way as to formulate sentences; to apply it to items in the world, employing as norms the various samples and training used in its explanation; and to hold oneself ultimately responsible to the use which other speakers make of it. This capacity presumably has some categorical basis in the mind, or more probably the brain, of the person whose capacity it is. We might think of this as a sort of mechanism. Now, it is surely reasonable to think that there are determinate truths about how this mechanism would react in regions of space and time that the person is incapable of reaching. For that mechanism, being part of the natural world, will be governed by causal laws whose operation is independent of space and time.¹⁴ In which case it will be determinate whether I would be inclined to describe anything in my surroundings as 'living' were I positioned in the Andromeda galaxy, or were I situated on Earth in the year 64 million BC. (Strictly: 'Whether someone with the same conceptual capacities as me would be inclined to describe . . .') This seems sufficient to imply that sentences such as 'There is life in the Andromeda galaxy' and 'There was life on Earth in 64 million BC' have determinate (though verification-transcendent) truth-values.

The account sketched here of the objectivity of contingent truth faces a number of problems and difficulties.¹⁵ For example, someone might con-

cede that it is determinate what would in fact be said were my capacity to be activated in remote regions of space or time, but object that this is insufficient to give the desired conclusion. For it needs also to be determinate that it would be *correct* for someone to say that. Yet does this not require that a norm of description can reach out to the world independently of my actual dispositions, which is just a return to logical objectivism? But this is to forget that the capacity in question is itself normative, in that it consists not only of dispositions to describe, but also of dispositions to correct and accept correction from others.¹⁶ So what counts as the correct exercise of my capacity is not something independent of it (except in so far as it partly depends upon the capacities of other speakers), but is rather one ingredient within it.

Another difficulty is that there may be many circumstances in which the brain-mechanism constitutive of the capacity would fail to respond in such a way as to record truth. For example, if situated on the surface of the sun or in deep space it would vaporise. But the reply to this is straightforward. It is that the counterfactuals about the mechanism's operation must include the proviso that it be insulated from anything which would interfere with its normal functioning (that is, that the person instantiating the mechanism should be protected by a heat-proof suit, and so on), where the idea of 'normal operation' is defined relative to the judgements people make in normal circumstances here on Earth. Then on this account 'There is life somewhere in the Andromeda galaxy' comes out as objectively true or false, despite the facts (a) that we are incapable of getting there, and (b) that in most places in that galaxy we would be incapable of existing unprotected, because there are determinate facts about how the mechanisms constitutive of our understanding of 'living creature' would respond if positioned at any given point in that galaxy, provided that it were protected in such a way as to ensure that it operates as it does here on Earth.

A third problem is that there are cases in which realising the antecedent of the counterfactual would constitute a material change in the subject-matter of the judgement. For example, consider the statement 'There are no intelligent agents in the Andromeda galaxy.' Since the mechanisms constitutive of understanding would have to be instantiated in an intelligent agent, they would, if positioned in the Andromeda galaxy, render it false rather than true that the galaxy contains no intelligent agents. The obvious response is to require that in examples of this sort the consequent of the counterfactual be adjusted accordingly. In this case what confers objectivity is that if an agent with my conceptual capacities were positioned

anywhere in the Andromeda galaxy, they would judge that there is no intelligent agent present *with the exception of themselves*.

Perhaps the most serious difficulty for the present defence of objectivity concerns theoretical statements in science. For in connection with a statement about sub-atomic particles, for instance, the idea of how the categorical bases of my conceptual capacities would respond if suitably placed can obviously get no grip. Now, one sort of response to this would be to point out that the objectivity of scientific truth is generally thought to be problematic in any case, quite apart from worries about the logical objectivist conception of meaning-to-world relations. We might then confine our account of objective truth to those domains where our pre-philosophical commitment to objectivity is strongest, namely truths about 'middle-sized' objects and processes, irrespective of their positions in space and time.

In fact it might be possible for us to make a more positive response, if it could be shown that there are limits to the thesis that scientific theories are underdetermined by their evidence. For suppose that the most that can be established is that there will always be more than one theory equally consistent with the data currently available to scientists, and that it would remain reasonable to believe that there would always be something to choose between theories assessed in the light of whatever criteria are accepted by scientists in good-making features of theories (simplicity, fecundity, etc.), in the face of all possible data. Then the existence of determinate truths about how the brain-mechanisms underlying a scientist's grasp of the concept *best theory* would react in the face of each item from the set of all possible data would be sufficient to confer objectivity on the claim that a given theory is better than any alternative.

The suggestion, then, is that my capacities and the state of the world may be said to settle all contingent truth-values between them, not independently of anything empirical (as logical objectivism requires), but in accordance with the laws of nature governing the operation of the mechanisms forming the categorical bases of the capacities. Since it is part of our ordinary world-view (of what we have reason to believe) that the universe consists of a wide range of mind-independent objects and events, some of which may be positioned beyond our cognitive access, and since it is also part of that world-view that capacities have categorical bases governed by laws which operate independently of spatial and temporal position, it then follows that we are committed to the objectivity of truth, believing all well-understood sentences to have determinate truth-values, quite apart from any commitment to logical objectivism.¹⁷

What, then, of the objectivity of conceptually necessary truth? Can this too survive the rejection of logical objectivism? In this case the situation is quite different; the argument to show it is presented best in the form of a dilemma. Firstly, suppose we are ontological realists about possible worlds. Then truths about all possible worlds will have to obtain in virtue of the operation of our conceptual capacities *within* each one of them. For our account of how a thought reaches out to another possible world to determine a truth value must surely be of the same kind as our account of how thought reaches out to different regions of the actual world: it will do so in virtue of the operation of our brain mechanisms in accordance with the laws of nature. But in that case it will follow that no statements are necessary, since the laws of nature vary across worlds. (In some worlds the laws of nature will be such that the mechanisms which are in fact constitutive of our understanding of 'red' and of 'coloured' respond in such a way that something can activate the first without activating the second.) So what emerges is that an ontological realist about possible worlds must be a logical objectivist about meaning-to-world relations, on pain of undermining one of the notions (conceptual necessity) that possible worlds were posited to explain. Or to put the point another way: in attacking logical objectivism we should have undermined the very notion of necessary truth, if we remained committed to the real existence of possible worlds.

Then suppose on the other hand that we are not ontological realists about possible worlds, regarding such talk as merely reflecting and 'objectifying' our discourse about what is and is not logically possible. In that case the closest that an appeal to the categorical bases of our conceptual capacities will get to vindicating the objectivity of a truth such as 'Necessarily anything red is coloured' is that the mechanisms constitutive of our understanding of 'red' and of 'coloured' are such that anything in the real world which activates the first would also activate the second. But there is nothing in this to distinguish the status of such a truth from a statement of natural necessity such as 'All mammals have hearts.' So if we have rejected logical objectivism, then the status as necessary of 'Anything red is coloured' must derive from something outside of it. For that status cannot be delivered by the idea of brain mechanisms operating in accordance with natural law alone. In which case, if 'Something is red but not coloured' is excluded from the domain, not only of actual but of possible truth, then this can only be because we have conventionally determined that it should be so.

If we deny both logical objectivism and ontological realism about possible worlds, then the necessity of 'Anything red is coloured' is best

explained as follows.¹⁸ Having acquired the capacities which constitute the understanding of 'red' and of 'coloured' respectively, and reflecting on our pattern of application of those terms, we realise that anything in the actual world which we should call red we would also call coloured. We also realise that we are unable, employing these capacities (and so in effect holding constant the causal laws which in fact govern their operation) to imagine circumstances where we should call an object red but be disinclined to describe it as coloured. As a result we elevate, by convention, the statement 'Anything red is coloured' to the status of necessity, using it from then on as an additional norm of description. We shall then say, if we are asked to consider worlds like the one imagined above where differences in the laws of nature mean that our capacities would operate differently, that in that case we should mean something different by the terms involved. But the status as such of a norm of description is not an objective, content-reflecting, feature of it, since someone could coherently accept that anything red will in fact be coloured (this is the most that is delivered by the underlying mechanisms) while remaining agnostic about its necessity. Conceptual necessity is thus not objective (mind-independent), but results from a combination of the limits of our imagination with a general convention requiring us to elevate the results of these limitations to the status of rules.

Either way, then (whether we are ontological realists about possible worlds or not), there is no place for the objectivity of necessary truth, given that the attack upon logical objectivism has been successful.

Summary

We have undermined the logical objectivism of *TLP*, using arguments gleaned from the later Wittgenstein's discussion of rule-following. The resulting identification of understanding with a practical capacity allows us to retain our belief in the objectivity of contingent truth, but requires us to give up our belief in the objectivity of conceptually necessary truth, embracing instead some form of radical conventionalism.¹⁹