

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

14 On Knowing One's Own Language¹

¹ The writing of this paper was made possible by a College Research Grant from Birkbeck College, University of London, and a visiting position in the Department of Philosophy at the University of California at Berkeley. Thanks are due to a number of people with whom I have had invaluable discussions about the material presented here. My thanks, in particular, go to Paul Boghossian, Marcia Cavell, Donald Davidson, Ariela Lazar, John McDowell, and Crispin Wright. I am indebted to Jim Higginbotham for a very rewarding discussion before the conference and his replies to my paper at the conference and in this volume. I am grateful to Michael Morris for a written comment. And most of all, to Maria Muller-Smith go my love and thanks for astute readings of drafts, insightful advice, and strong support.

Barry C. Smith

We rely on language to know the minds of others, but does language have a role to play in knowing our own minds? To suppose it does is to look for a connection between mastery of a language and the epistemic relation we bear to our inner lives. What could such a connection consist in? To explore this, I shall examine strategies for explaining self-knowledge in terms of the use we make of language to express and report our mental states. Success in these strategies will depend on the view we take of speakers' understanding of the words they use to speak their minds. The key is to avoid circularity in the account of how they know what they mean; for if knowing what one is saying in speaking a language provides a means of knowing one's own mind, it cannot simply be a part of it. I shall look at ways in which we might proceed here, and examine whether the strategy can make room for a genuinely first-person point of view. But first let me try to motivate the problem of self-knowledge.

I The Problem of Self-Knowledge

It is hard at first to see what the problem of self-knowledge is supposed to be. As thinkers, we take ourselves to have immediate knowledge of

end p.391

our thoughts and feelings and to know our minds intimately and in a way that no one else can know them, all of which can easily incline us towards a conception of the mind as a repository for inner thoughts, a sphere whose contents are inspected privately. But however intuitive this picture seems from our own point of view, we conceive the mind quite differently when thinking about others. In taking up the third-person perspective, we are convinced that we know a lot about other people's minds. Not, to be sure, in the way we know our own; but we are convinced that others' behaviour can teach us a good deal about what they are thinking, what they want, and what they are up to. We see others' minds at work in their acts and utterances, and have little difficulty in describing their activity as displaying their preferences, showing their feelings, and revealing their motives. Were the workings of people's minds not on show in this way, we would be unable to recognize them as persons.

It would appear, then, that common sense presents us with conflicting views on the nature of mental life, depending on whether we are thinking of the accessibility of minds to others or of their accessibility to themselves. We are left in the curious position of believing that whereas the minds of others can be displayed in observable behaviour, the contents of our own minds are a matter for private consumption. Clearly, this is nothing more than a comforting illusion. For we are just as likely as anyone else to reveal aspects of our mental lives publicly.

The problem arises when we try to square this objective and outward aspect of the mental with the special way in which we know our own minds from the first-person perspective. How can states whose natures belong partly in the public sphere be so readily available to us from the first-person point of view? I take this to be the key problem raised by self-knowledge.

As subjects, we are inclined to think of such states as being automatically available to us, part of our inner lives. But this merely makes the philosophical difficulties more acute. How can the facts of mental life be part of the inner world of a subject if they are also objective facts knowable by others on the basis of outwardly observable behaviour?

The problem we confront here partly stems from a lingering attachment to the intuitive picture of the inner with which we began. That picture took our states of mind to be inner in a wholly private and subjective sense. But the mental states that make up our inner lives can be known by others. We need a new conception that does not remove the mental from other people's scrutinizing gaze. After all, it is not privacy that makes for an inner life. It is the fact that we have unmediated knowledge of

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

end p.392

our mental states. Each of us is familiar with the details of inner life through the thoughts and feelings that are immediately known to us. Others do not have this unmediated knowledge of our minds, although they will know their own minds in this way. So it is not the privacy of the object known, but the exclusive means of knowing it, that makes room for the inner world. We credit minded creatures with an inner life, or point of view, only if they have the capacity to know their own minds in this way. The inner and public aspects of the mental are connected.

Still, the way we know and the way others know what we are thinking differ markedly. We know our own minds, when we do, in a direct and authoritative manner. We do not have to observe our own behaviour or consult the opinion of others to find out what we think and feel. We just know. We can offer no grounds for what we take ourselves to be thinking; nor are we expected to. Others do not ask us for justification, because they regard us as authorities on matters of our own psychology. We just assume that we know our own minds, and it makes little sense, in general, for us to doubt this. Hence, far from undermining our claim to knowledge, the absence of grounds, in the case of claims about our current psychological states, gives us an authority that our ascriptions of psychological states to others, and their ascriptions to us, simply lack. In this way, first-person authority is the hallmark of self-knowledge, and provides the most telling asymmetry between first- and third-person knowledge of someone's mind. The problem is to find a way to do justice to the special authoritative character of self-knowledge, while respecting the publicly determinable character of the psychological states known. This is the problem of self-knowledge.

The problem tends to get overlooked if, like Descartes, we focus on sceptical questions. For we cannot doubt that we have knowledge of our own minds. Taking this for granted, it appears to be knowledge of other minds and of the external world which we need to account for. However, the ease of acquiring self-knowledge, compared to the efforts it takes to figure out what others are thinking, should not mislead us as to the nature of that knowledge or the ease, or otherwise, of explaining it. Just assuring myself that I know *that* I am thinking does nothing to illuminate how I know *what* I am thinking, or to settle questions about the relation of self-knowledge to knowledge of other minds or to the world. So we must ask: How is it possible to know what we are thinking?, where this is not a sceptical question requiring an answer that will assure us that such knowledge really is possible. That is scarcely coherent. Rather, we are asking for an explanation: How do we know what we are thinking, given that we usually do?

end p.393

Traditional epistemologies of mind are of little help, since they often end up at odds with elements of common sense we are keen to preserve. Cartesians, for example, conceive the domain of the mental as a private realm whose layout is fully transparent to its subject but inaccessible to others. Mental states are constituted by how things appear to the subject and are answerable to nothing else. This appears to ensure the subject knowledge of what he is thinking, but it puts at risk his knowledge of other minds and access to the world. Moreover, it utterly distorts the nature of self-knowledge. For we are neither infallible nor incorrigible judges of our own intentional states. Our actual epistemic predicament is at once less certain and more difficult to explain. When we believe, desire, or intend something, we usually know that we do. But this is not always the case. And since what we believe, desire, or intend is open to determination by others on the basis of our behaviour, our states of mind must be accessible from perspectives other than our own. It is this knowledge that others can have of us which makes lapses in our self-knowledge intelligible. For their knowledge of our minds need not be arrived at via the views we have of what we are thinking. Our behaviour can reveal a good deal more about our outlook and attitudes than we suppose. And this is likely to be the case even at times when, for whatever reason, we are blind to our own intentions. Thus people, if they are sufficiently insightful, may draw our attention to aspects of our psychology about which we are deceived or un-self-knowing. They may convince us that our intentions are less than honourable, and we may feel the shame of motives late revealed. In this way other people's perceptions can provide a valuable source of self-knowledge: an insight lost on the Cartesian.

However, it would be wrong to conclude with the behaviourist that knowledge of mind is based solely on what is publicly observable. I usually know what I am thinking without consulting the evidence available to others. And it is this more familiar way of knowing our own minds that we need to account for. The behaviourist is simply at a loss to explain how a person can form correct opinions about her states of mind when she doesn't rely on behavioural observations to do so. Having nothing more to go on than what is publicly observable, the behaviourist is incapable of recognizing the peculiar immediacy of our self-knowledge. At best, he will see the advantage that subjects enjoy in reporting their mental states as a function of the enhanced opportunities they have to witness their own activities and infer conclusions about their mental lives. But we need make no such inferences to secure first-personal advantage. We are peculiarly qualified to pronounce on our own states of

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

mind, and we do this without reference to our outwardly observable activities.

end p.394

The difficulty is to explain how this is possible when knowledge of our mental states is not a matter of each of us having privileged access to some private realm. To its credit, behaviourism recognizes the public aspect of mental states, but it loses sight of the inner, and with it the subject's point of view.

It should now be clear why neither Cartesian nor behaviourist conceptions of mind can explain self-knowledge. Each fails to recognize that it is in the nature of mental states—particularly intentional states like belief, desire, and intention—to be known both first-personally and third-personally. Both the Cartesian and the behaviourist, in their different ways, recommend a conception according to which the facts of mental life are exhausted from a single perspective. Cartesians assume that mental states are always and only available from the first-person point of view, while behaviourists suppose that they are exclusively available third-personally. Each stresses an important aspect of the mind at the cost of neglecting a feature stressed by the other. The behaviourist is right to point out that we often know the minds of others on the basis of observing what they say and do, but is wrong to suppose that self-knowledge is continuous with knowledge of other minds. The Cartesian is right to stress that we can be directly aware of what we are thinking, but wrong to suppose that this is a matter of acquaintance with a private inner item. The special unmediated knowledge we have of our own minds furnishes us with an inner life, but this does not preclude others from knowing the states of mind we are in. Other people's inner lives are not always hidden from us, nor ours from them.²

² Wittgenstein's remarks here are useful reminders of how we should guard against faulty conceptions of the inner: 'One can say "He is hiding his feelings". But that means that it is not a priori that they are always hidden. Or: there are two statements contradicting one another: one is that feelings are essentially hidden; the other that someone is hiding his feelings from me' (1992: 35e).

Nevertheless, each individual is uniquely placed to judge what he, or she, wants, intends, or believes. And we need to know what puts one in such a position. What makes one groundlessly authoritative about attitudes that are accessible from perspectives other than one's own?

II The Nature of Self-Knowledge

Initially it might appear that the public side to mental life threatens the possibility of self-knowledge. But on the contrary, publicity ensures that self-knowledge lays claim to a genuine subject-matter, the external

end p.395

dimension ensuring that whether a subject is in a given mental state is an objective, or, at any rate, intersubjective, matter answerable to more than just the subject's opinions. Thus, the states that are available to us as part of our inner world can be credited with an objective reality over and above our immediate impressions of them. What is puzzling is just how one's immediate impressions can provide one with knowledge of objective empirical facts—one's being in certain publicly determinable states of mind. Why should the immediate, groundless psychological claims we make about ourselves be largely correct? Having groundless authority about what is objectively the case is what makes this variety of knowledge so puzzling. How are we to explain it?

Our mental states are not the sorts of things to which we are guaranteed access: *being in a mental state* is one thing; *knowing which mental states one is in* is another. Something is required of us to close this epistemic gap. Yet, our thoughts and feelings are so familiar to us that it is hard to consider our knowledge of them as amounting to a genuine cognitive accomplishment. We tend to think our thoughts are available to us in the very act of thinking them. This is because no gap opens up, *from the first-person point of view*, between being in a mental state and knowing one is. The thinker will be unaware of the potential for a gap until others draw his attention to it, or he appreciates such a gap in them. Whether one is in a mental state or not is not a matter exclusively settled by how things appear to the person in that state. As we have noted, the existence of an independent third-person perspective secures a certain objectivity for claims about a person's psychology. It is from these other-person perspectives that one can establish the distinction between *how things seem* to a person psychologically and *how things are*: the distinction necessary to guarantee the objectivity, or, at any rate, intersubjectivity, of a subject's claims. Appreciating this requires that one recognize that one's inner world is part of the world of others. Without this, 'no distinction between "seems right" and "is right" opens up from the subject's point of view' (McDowell 1991: 160).

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

What we need to do now is to connect up these two distinctions: the distinction between *being in a mental state* and *knowing that one is* and the distinction between *how things seem to a thinker*, psychologically, and *how they are*. Obviously, being in a psychological state will correspond to how things are, but we have yet to understand why, and under what conditions, knowing what psychological state we are in will correspond to how things seem to us. Why does how things seem to us amount to knowing what psychological states we are in?

In the psychological domain we cannot separate the nature of mental states from their epistemology. Our fix on how things are with a thinker—the external check on his impressions—is determined by the

end p.396

psychological states we would be properly justified in attributing to him on the basis of what he says and does, against a vast background of previously ascribed psychological states that make sense of his behaviour to date. Thus, publicity is guaranteed in our conception of these objective states of affairs. As Davidson puts it, states like belief and desire

are just those states whose contents can be discovered in well-known ways. If other people or creatures are in states not discoverable by these methods, it can be, not because the methods fail us, but because these states are not correctly called states of mind—they are not beliefs, desires, wishes or intentions. (Davidson 1989: 160)

What we now need to discover is whether, and if so, how, a subject's unsupported views about what he currently believes, wants, or intends can count amongst these well-known ways. Why should a subject's merely *thinking* that he is in a certain mental state be a reliable guide to his *actual* mental state? More is needed than his merely supposing that he has these beliefs, desires, or intentions, to make his suppositions true. The third-personal grounds for ascription provide an independent basis for evaluation. So, we can say that a subject counts as knowing his own mind provided he comes up with ascriptions of attitudes to himself which others would be prepared to ascribe to him given all the available evidence and in accordance with the correct standards for mental attribution. In this way, claims to self-knowledge, although not based on behavioural evidence, will still be answerable to it. For if a subject cannot be correctly said to believe, desire, or intend something according to well-known methods of ascription, his merely claiming that he does will establish nothing.³

³ Hence more is needed than the suggestion by Rorty (1981) that it is just a courtesy of the interpreter to grant subjects authority for what they say about their own states of mind. We do not slacken all the standards we exercise in assessing what people tell us just because, in this case, the subject-matter concerns their states of mind. They are authoritative, but the authority can be defeasible on cognitive grounds, and not merely lack of courtesy.

We can count a subject's current impression of what he is thinking as one amongst other well-known ways of discovering his present state of mind just so long as this way of knowing can be harmonized with the other methods of discovery. This provides an important publicity constraint on any adequate account of self-knowledge. It means that the problem of self-knowledge and the problem of other minds arise with respect to the same state of a thinker's mind. We must solve them simultaneously to produce a complete, satisfying epistemology of the mental.

The difficulty is to see why, of the two ways of discovering what someone is thinking, the first-personal means is authoritative even though

end p.397

someone's first-personal claims about what he is thinking are ultimately answerable for their truth to third-personal claims about which attitudes it would be correct to ascribe to him. This points to a certain strain between first-person authority and the objectivity of self-knowledge. The more subjective a thinker's judgement, the less it concerns anything other than how things seem to the subject. And in such cases, it is easy to see why the thinker should be an authority about this. By contrast, when there is an objective fact of the matter about which the subject can be right or wrong, it is hard to see why, without having any grounds for his judgements, he should be, in Crispin Wright's words, effortlessly authoritative about the facts.

There is a prima-facie tension between first-person authority and objectivity. Objectivity requires that there be room for a distinction between how things seem to the subject and how they are, between whether a judgement seems right or is right. While having first-person authority is based on there being no appreciable difference between how things seem and how they are. I am authoritative just in case, typically, whatever seems right to me is right. First-person authority requires us to close the gap between having beliefs and having opinions about them, while objectivity insists on leaving room for a gap between them. However, first-person authority is not really incompatible with the objectivity of self-knowledge. There is room for a distinction between how things seem to the subject and how things are, even if they coincide, because the existence of

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

perspectives other than the subject's own provide us with an independent means of knowing how things are with him. We use these to confirm, or challenge, the subject's self-assessments. Meanwhile, the fact that how things seem to a subject, psychologically, is usually how they are, means that he retains authority about his current psychological states. What remains to be explained is why being in a mental state, and thinking that one is, should so regularly coincide. Why should one's groundlessly authoritative conviction that one is in a particular mental state be a good guide to the independently confirmable fact of the matter?

The problem of self-knowledge has become a little clearer. We asked what closes the epistemic gap between being in a mental state and knowing that one is. The answer we gave has two parts: (i) without any grounds subjects simply take themselves to be in particular mental states; and (ii) whenever they take themselves to be in particular mental states, they are usually right. Therefore, any solution to the problem will have to address two issues. What equips thinkers to make psychological judgements about themselves, and why do these judgements mostly relate

end p.398

thinkers to their inner lives? A solution to the problem depends on finding satisfactory answers to these key questions:

- (1) What is it for things to seem a certain way to a person, psychologically?
- (2) Why is it that how things seem, psychologically, usually provides knowledge of how they are?

We need an answer to (1) to make room for the subjective point of view of a thinker, and an answer to (2) to make good the thinker's claim to objective knowledge of his own mental states. No treatment of self-knowledge will be satisfactory unless it addresses both of these requirements. For only then can we reconcile the inner and outer aspects of mind: how things subjectively seem to us and how they objectively are.

III Models of Self-Knowledge

What makes objective psychological facts about ourselves subjectively available to us? This question presents no difficulties for the Cartesian, for whom there is only what is subjectively available to us; or, as Cartesians would have it, what is subjectively available always coincides with the facts. The motivations for this view are clear, but deeply flawed. As we have seen, no gap opens up from the subject's point of view between what is, and what seems to be, going on in the subject's mind. And since, for the Cartesian with regard to the mental, the subject is incorrigible—his view of what is going through his mind is the only view there is—no gap can open up *at all* between how things seem to the subject and how they are. The assumption is that the mental landscape is so well lit that there are no parts of it he cannot see. This makes him an infallible judge of his own mind; and infallibility (more than) guarantees first-person authority. There is no gap to close between what we believe and our opinions about this, so we need no justification for our view of what we are thinking. And since there is no justification for the sceptic to call into question, this makes self-attributions immune to sceptical doubt and the firmest foundation for all knowledge.

However, this is a Pyrrhic victory over the sceptic, since it is at the cost of surrendering the objectivity of what we judge. If there is only how things seem to us, then whatever seems right, is right, and as Wittgenstein tells us, this just means we cannot speak of right here. The Cartesian looks as if he is setting out to answer question (2) by telling us that how things seem will always coincide with how they are in the

end p.399

mind; but in denying the publicity requirement—the possibility of another knowing what states of mind we are in—he cannot secure any objectivity for our self-ascriptions.

It is also worth nothing that Cartesianism gives us no help with (1). It simply takes for granted the subjective point of view, with occasional hints about the mind's eye perceiving its own inner items. It is not hard to see why the perceptual model seems attractive. Perception can give us immediate knowledge of objective states of affairs (on at least some views), so the 'inner' perception model may still appeal as an explanation of how we know our own minds even though we have rejected Cartesianism. The rejection of the Cartesian picture obliges us to acknowledge that what is inwardly perceived is not a private inner item. Others can have knowledge of it too.

What are the prospects for such an inner perception model of self-knowledge? How might it provide answers to (1) and (2) above? With respect to (2), to secure the claim to knowledge of an objective psychological state of affairs, it is necessary to show that the states whose contents are presented to us directly, through their perceptible features, are the very same states we would be apt to ascribe on third-personal grounds, without reference to those features. Only then can we be sure

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

that what is inwardly observable by the subject is not solely inner and private, but discoverable by others in well-known ways. While in relation to (1), the thought is that since others lack direct *perceptual* access to the contents of my mind, they cannot be presented with my states in the way I am, and will lack the authority with which I judge what I am thinking—hence the asymmetry between the first- and third-person perspectives.

There is no disguising the fact that the non-Cartesian dimensions to the mental make trouble for the inner perception model. We do not have exclusive access to our intentional states, so there must be more to the mental than meets the mind's eye. And since we are not infallible judges of our inner lives, inner vision, like outer vision, may be less than twenty-twenty. Given these limitations, can it be guaranteed that what we perceive inwardly are the same states that are individuated third-personally, non-perceptually? To turn the trick, what we inwardly observe has to be something that spans the first- and third-person divide, something that can be individuated on purely third-personal grounds. But although we have immediate perceptual access to our states, we do not have such access to their individuating conditions. This creates a problem, because public and external factors can enter into the identification of the states which people ascribe to us. In ascribing particular thoughts to a thinker, people observe not only the thinker's behaviour

end p.400

but his connections to things in his environment. According to externalists about content, these factors play an essential role in individuating many of our mental states. But if the contents of many of our thoughts are determined in part by our causal or socially mediated relations to the people, places, or things in our vicinity, there will be public aspects of our thinking more easily appreciated by those who can see exactly how we are placed in relation to these objects. So how do we retain first-person authority? Whatever *is* directly apprehended in inner perception may fail to discriminate between environmentally distinct states with internally *similar* perceptible properties.⁴

⁴ Notice that failure to discriminate between thoughts will not prevent it being the case, on each occasion, that we are in inner perceptual contact with the particular thought actually tokened, *whatever thought that is*. This may not satisfy those who had hoped for a more substantive account of first-person authority.

Externalism is thought to present a special problem for first-person authority, but in fact it is just part of the general problem of selfknowledge outlined above. It adds a further refinement to the question of how we are to reconcile the inner with the public aspects of the mind.⁵

⁵ Notice that no advantage is gained by being an individualist about the mental. Those who suppose that it is scientific psychology that individuates our internal states still need to provide an account of how a person knows what internal states he is in without a theory, while those who favour a purely phenomenological account of individualism have still to explain how such states can have any objective content. The debate between externalists and individualists is simply orthogonal to the issue of self-knowledge.

But is it compatible with the inner perception model of selfknowledge? The only way we can know what externally individuated states of mind we are in is by perceiving the contents of those states. So now we need a detailed account of how the externally individuated contents of our states are perceptually presented to the mind, and, as yet, we do not have one. The trouble is that the evidential bases on which other people ascribe attitudes to us are only contingently connected to these immediately apprehensible features of our thoughts. No explanation is offered of the coincidence in properties here. Why should there be two quite different ways of being presented with the *content* of a mental state? It remains wholly unexplained why the content of a state should exhibit perceptible features *and* bear connections to the subject's publicly displayed behaviour. This appears more like a re-description of the problem than a solution to it.

There are other difficulties too. Intentional states can occur without registering their presence in the mind. We know this because there can be good grounds for ascribing states to someone who shows no inner recognition of them. Hence the properties by which we consciously apprehend mental states (or our conscious apprehension of them) cannot be

end p.401

essential to those states. Why, then, should we suppose that the directly apprehensible properties of mental states are good guides to their true natures? And, if we can be mistaken about the states we are in, not being infallible, then how things appear to us, psychologically, can fail to show us how things are. The inner appearances and the mental states can come apart. The inwardly perceptible features of our states will not always coincide with the mental states they purport to reveal. Does this mean that intentional properties can appear to be before the mind without an appropriate intentional state being present? Or does it mean that inner states can take on different appearances, as if there were tricks of the light? Can we misperceive the contents of the thought we are thinking? A more likely explanation is that we can *misjudge* our current states

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

of mind, as when we misjudge a mood, or a feeling. This would make it the element of judgement, not perception, that we should concentrate on. The inner appearances do not always play a verifying role, nor do they explain our mistaken views of our own psychology. Inner perception may have some claim to cast light on the epistemological asymmetry between first- and third-person knowledge, but it has some way to go to explain how there can be a common object of knowledge.⁶

⁶ For attempts to deal with these and other objections see Macdonald, Ch. 5 above.

Whether or not these problems can be ironed out, the most telling objection to the inner perception model is the absence of any phenomenal or perceptual quality to accompany many of the beliefs and intentions over which we have first-person authority. In many cases there is simply no perceptible item before the mind from which we can read off our current thoughts or motives. In the case of intentions, as Moran (1988) points out, discovering what one intends is more akin to making up one's mind than reading off the properties of a perceptually salient inner item. And what distinctive perceptual experience corresponds to the intention to read more contemporary fiction or to go to bed earlier? For many of our psychological states, there is simply no experiential item, but only the judgements we make about them. There may be nothing at which to direct my inner gaze—whatever that means—and yet I can be firmly of the opinion that I want to go to Scotland. I have no grounds for the judgement that this is what I want; *a fortiori*, I have no perceptual grounds for taking that to be what I want. In coming to have knowledge of our own minds, we arrive at opinions, without inference or evidence, about what we are thinking, and we are usually right. To explain this, we need a radically different approach based on a non-perceptual epistemology.

What are the options here? Why are we so effortlessly authoritative about our present states of mind? An answer, oft repeated, is that it

end p.402

is a fact about us that when we are in first-order mental states with a certain content, this reliably causes the tokening of second-order states whose contents are about, or contain, the content of the first-order states. Thus the capacity for self-knowledge is thought to consist in the presence of a reliable connection between first-order and higher-order mental states. But where a gap exists, like the one pointed out above, between *being in a mental state* and *knowing that one is*, it is hard to see why *being in a another state* (albeit reliably caused by the first, and related to it in content) amounts to knowing the first state one is in. How does *being* in the second-order state enable us to *know* its content? Accounts of the reliable causal connection between second- and first-order states are simply silent on the crucial epistemic gap. So I shall not pursue the reliabilist option any further.⁷

⁷ Tyler Burge (1988) has proposed a non-causal way of relating first- and second-order thoughts in the case of self-verifying claims of the form 'I am thinking the thought that water is wet'. In such cases the external factors that determine the embedded content contribute to determining the content of the reflective thought. In this case, just being able to think these thoughts ensures comprehension of what one is thinking, and the tokening of the thought one is thinking about. However, not all cases of self-knowledge take this form.

Some may be tempted to say that we know our thoughts and feelings by *introspection*. This is undoubtedly true, if by 'introspection' we mean no more than having direct knowledge of what we think and feel. But this takes us no further forward. On the other hand, if we mean by 'introspection' some special and privileged relation between a thinker and his thoughts, what is this relation? It cannot be perceptual. But we have nothing more to go on. Talk of introspection appears to be no more than a place-holder for some, as yet, unspecified account.

Another popular answer is that consciousness explains self-knowledge. A subject knows what mental states he is in when he is *conscious of* them. Once again, if we mean by 'being conscious of' no more than 'having immediate knowledge of' or 'having access to', we are no further forward. On the other hand, if consciousness plays a more substantive role, we are owed an explanation of how it enables a thinker to know *which* states of mind he is in. How do momentary episodes in conscious experience constitute knowledge of beliefs and intentions which sustain rational connections to other attitudes and to actions. Conscious experiences *per se* are too transitory to provide what is needed here. No such ephemeral items could *by themselves* sustain the rational connections to behaviour, and further mental states, necessary to individuate the content of thought. Only the exercise of conceptual judgement could do this. Even in the case of our knowledge of our own

end p.403

conscious experiences, it is not merely their being conscious that gives us knowledge, but our having the classificatory competence to identify and discriminate them. No one denies that Wittgenstein's private linguist is conscious. What is at

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

issue is whether he has the wherewithal to set up and maintain objective standards for judgements regarding his own experiences.

None of the foregoing epistemologies successfully reconciles the immediate availability of thoughts to a thinker with the public and external dimension of thought. Observational epistemologies force us to look either inside for some special subjective item or outside at patterns in our behaviour. But mental states should be conceived as neither inwardly nor outwardly perceived items. Perception cannot locate states with both an inner and an outer dimension properly. We need a fundamentally different epistemology for our psychological states. And I want now to consider whether the alternative epistemology should be based on our use of language to identify and discriminate our psychological states.

IV The Linguistic Strategy

A language-based epistemology of mind will have to offer a solution to the problem of self-knowledge outlined above. It must explain what equips someone to make groundless but authoritative judgements about himself, as well as explain what makes his judgements mostly correct. Remember that these judgements will be correct when how the subject takes things to be is how they are. This is a matter of a thinker's being in particular mental states, an objective matter of his meeting the third-personally accessible conditions for being ascribed such states. In this way, publicity is guaranteed. Any mental states we know ourselves to be in must be knowable (not necessarily known) by others. So any satisfactory solution must accommodate the fact that the same mental states will be accessible from the first- and the third-person points of view. I will also have to preserve the asymmetry due to first-person authority.

Now if language has a key role to play in how others know what we are thinking, and the same states of mind are accessible from both perspectives, could it have a similar role to play in how we know our own minds?

The linguistic strategy has a lot to recommend it. Meanings, like mental states, have a first- and a third-person epistemology. There are also externalist elements to linguistic meaning. (Putnam's celebrated argument that meanings are not in the head was the first defence of externalism.)

end p.404

In so far as we convey meaning to others, the words by which we make our thoughts available to them are words which express those thoughts for us. Hence the same thing can be apprehended from the first- and the third-person point of view, although there is also the question of authority. You may have to work out what my words mean, whereas I do not. The meaning of my words is immediately available to me. But this is not to deny the public nature of meaning. I cannot make my words mean whatever I want them to; nor do mere subjective impressions of meaning constitute knowledge of meaning. Whether or not I use words as others do, the publicity requirement ensures that all I can mean by my words is all I can be known to mean. There is still asymmetry, because I have *immediate* knowledge of what I mean. Clearly, there are interesting parallels between the linguistic and the mental, but more is needed if we are to use language to explain self-knowledge.

Language can enter the picture in different ways, and there are stronger or weaker positions on the connection between knowing one's language and knowing one's own mind, depending on whether one's knowledge of language is considered to be constitutive of, necessary, or merely sufficient for knowing one's mind. The choice here depends on one's view of the precise relations between thought and language. At one end of the spectrum is the view that language is the vehicle for thought. This is a view held by John McDowell, amongst others. He tells us that 'acquiring a first language is, not learning a behavioural outlet for antecedent states of mind, but becoming *minded* in ways that language is anyway able to express'. For McDowell, this is because 'our ability to have dealings with content can be, not a mere natural endowment . . . but an achievement, which an individual attains by acquiring membership of a linguistic community' (1987: 74). On this view, language is essential for thought, and knowledge of language plays a constitutive role in knowing one's mental states. I know what I am thinking by understanding the words that furnish me with thought.

An attractive feature of this view is that it avoids making our ordinary knowledge of what we are thinking too reflective. Finding the words to express thoughts is what it is to think, and our immediate comprehension of our words gives us immediate knowledge of what we are thinking. If understanding is present in the act of saying something, we shall seldom need to make second-order judgements to know about the intentional states we are in. Our comprehending use of the meaning and form of language makes conceptually structured thought possible. We know our beliefs through what we are prepared to assert. (Of course, this is the case only if understanding embraces knowledge of the force of our utterances, as it does for McDowell.) I recognize the content of

end p.405

what I am saying because I understand the words I speak.⁸

⁸ It may be objected that we do not always have to speak our thoughts aloud. (These days many people feel they do!) Nevertheless, most of us are all too familiar with the restless interior monologue.

But what is it for words to have the meanings they have for us? We need an answer to this question if linguistic understanding is to serve in the explanation of self-knowledge. What equips a speaker to take his words to have the meanings they do? Notice also that, in order to secure first-person authority, it will have to be shown that we know what our words mean in a way that differs from how others know this.⁹

⁹ I suspect McDowell would balk at this, though it is worth noting that the idea finds some endorsement in Wittgenstein: 'My own relation to my words is wholly different from other people's' (1953, iix: 192e). As he says, 'I do not listen to them and thereby learn something about myself' (1992: 9e). 'If I listened to the words of my mouth, I might say that someone else was speaking out of my mouth' (1953, iix: 192e)—a clear indication of why linguistic understanding cannot consist solely in a perceptual capacity, as McDowell sometimes appears to think. (Note that this would reintroduce a perceptual epistemology of mind. We would no longer have *inner* perception of the content of mental states, but we would be able to perceive the contents of mental states, our own and others', by perceiving the linguistic meanings of utterances that expressed them.) We need to focus instead on the states of linguistic knowledge at work in both the perception and the production of speech.

Having authoritative knowledge of meaning gives us authoritative knowledge of the thoughts we express. Reflection can serve up judgements about what we think (doubt, hope for, fear, or intend), though we will have an immediate and correct view of what we are thinking just so long as we understand what we are saying. An account of self-knowledge becomes an account of how I know what I mean by the words that express my thoughts.

Notice that such a strong position commits us to denying thought to animals and pre-linguistic infants: a highly controversial claim which requires detailed empirical as well as philosophical investigation. However, nothing so strong is required for one to exploit the linguistic strategy. For, whether or not thoughts are essentially linguistic, we often find out what we think by saying it. Language can help us to formulate our thoughts more precisely. And the greater our linguistic competence, the greater the clarity with which we realize what we really think. On this view, although it is not necessary, knowledge of language is sufficient on these occasions for knowing what one is thinking. It provides one means among others of knowing one's mind. I shall return to this view below.

There is another way that language can enter the picture. This is the view that language provides the folk-psychological concepts of belief,

end p.406

desire, and intention, with which we taxonomize our psychological states. The vocabulary of the attitudes, together with a potentially infinite repertoire of declarative sentences, embedded by constructions like 'X believes that', 'X intends that', 'X desires that', provide us with the means to identify and discriminate the thoughts of a rational agent.¹⁰

¹⁰ Again, there are weaker and stronger readings of this position. On the weaker version, language may not be necessary for thought, but it is necessary for *thought about thought*. So, while the objects of self-knowledge may not be linguistic, language and our knowledge of it will still play an essential role in the epistemology of mind. The stronger reading maintains that to have thoughts—in particular, beliefs—one must have the concept of belief, thus be able to think about thoughts. And since language is necessary for thought about thought, it is therefore necessary for thought, though not constitutive of it. Once again, this denies thought to language-less creatures.

And since the nature of mental states is bound up with our well-known means of discovering them, the same indispensable taxonomy is at work, so the story goes, in finely individuating our own thoughts. The story, in its most subtle and systematic form, is given to us by Donald Davidson. To see the full force of it, we have to begin with our knowledge of other minds.

All we see of other people is their behaviour; yet we often know what they are thinking. We arrive at knowledge of their minds by interpreting their behaviour, and we do this by making sense of their activities in rational terms. In order to see someone as a rational agent, we must be able to see his behaviour as action, and we do this by describing bits of behaviour in intentional terms. This brings in its wake attributions to the agent of beliefs and desires that would rationalize his actions. We cite as his reasons for acting, particular beliefs and desires that it would be reasonable for him to have, given what we describe him as doing, and given what else he thinks and wants, as judged against the other psychological ascriptions we would be prepared to make of him. Each attribution is tested by how well it fits into a larger scheme of interpretation of what he says and does elsewhere. This will be a coherent, holistic pattern of ascription of attitudes and meanings that makes the

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

best overall sense of the agent's actions and utterances. Only when an agent's behaviour can sustain these intentional descriptions and attributions can he be said to have beliefs, desires, and intentions. The intentional descriptions are our means of individuating his mental states, but they also play a constitutive role in identifying the nature of mental states. For Davidson, particular events in a person's physical history are, at the same time, events in his mental life. But an event is mental only as described. It is only when an otherwise physically describable event, or state, can be given an intentional description that it can be said to be

end p.407

a mental state at all. And since we rely on the language of belief–desire psychology to individuate mental states finely enough for the rational explanation of behaviour, so the taxonomy of folk-psychological terms will have a role to play in fixing the identity of mental states. According to this view, it is necessary to deploy the intentional vocabulary of psychology to know one's own mind. But so far, we have no explanation of a subject's ability to ascribe attitudes to himself; far less have we any explanation of why these groundless self-ascriptions should carry the authority they are standardly granted.

Davidson's is an essentially third-personal view of meaning and mind: the mental and the meaningful are constituted by the intentional categories we impose on behaviour to make sense of ourselves and others in rational terms. All there is to a thinker's mind is all that can be discovered through the method of interpretation.¹¹

¹¹ 'What a fully informed interpreter could learn about what a speaker means is all there is to learn; the same goes for what the speaker believes' (Davidson 1986a: 315).

A theory of interpretation is a device to show how we can provide a systematic justification for the ascription of belief, desire, intention, meaning, and action to a person who exhibits a suitably wide range of behaviours. To be correct, any claims a thinker makes about himself must be justifiable in terms of an overall interpretation of his behaviour.

Notice that it is not Davidson's intention to deny that thinkers have a point of view of their own. He accepts the asymmetry between first- and other-person ascriptions of mental states and the existence of first-person authority, which he understands as the presumption of correctness which attaches to first-person attributions. It is just that, ideally, attributions from the first- and third-person perspectives should converge. The phenomenon of convergence is reflected in the practice of reporting beliefs and other attitudes. When I say, 'I believe (desire, hope, fear) that the guests will be on time', this is true if and only if the corresponding statement, 'He believes (desires, hopes, fears) that the guests will be on time' is true, when said by you about me. There is an agreement in truth-conditions, but there is an asymmetry in how the truth of each claim is arrived at. The problem, of course, is how to reconcile the *truthconditional symmetry* and the *epistemological asymmetry* between self- and other-ascriptions to a person of a given mental state.¹²

¹² I owe this way of putting the problem to Marcia Cavell.

Because of the agreement here, we know that settling the truth-values of thirdperson claims will settle the truth-values of the corresponding firstperson claims. But that is not how subjects arrive at the truth of their

end p.408

self-ascriptions. On the contrary, when we attribute attitudes to people on the basis of what they say and do, we expect them to know these attitudes, and are surprised if they do not. We often take first-person pronouncements as confirmation of our views about what they are thinking; what people say about themselves chimes in with the views we had arrived at by previous knowledge of their speech and action. In this way, when we ascribe attitudes to people, we try to figure out *their view* of the situation. For unless the beliefs and desires we ascribe to them are the states that shape *their* outlook on the world and give *them* reasons to act, we shall not be really understanding their behaviour. Thus, it is the aim of any accurate interpretation to respect the subject's point of view. We want to ascribe to people the attitudes they really have and could confirm by what they say about themselves first-personally.¹³

¹³ Notice that a subject's being in a mental state is not enough to ensure that it contributes to his point of view. Merely to take it on someone else's word—that is, on good third-person grounds—that I have certain intentions or motives will not count either. The subjective outlook from which a person views the world is structured by states of mind about which he is authoritative. This determines the extent to which he is self-knowing.

This may go some way to assuaging a worry about our self-ascriptions having to be too strictly accountable to the judgements which others would be prepared to ascribe to us third-personally. Still, there remains the question, for Davidson, of how the person himself knows what he is thinking. We know the minds of others by interpreting their behaviour. But we

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

hardly ever need to study ourselves to come up with views about what we are thinking. The key question for Davidson is:

(I) How do we know our own minds given that we do not interpret ourselves?

V Linguistic Models of Self-Knowledge

The predicament is by now familiar. We seem to come up with claims about our own minds which are not based on evidence or inference, but which carry greater authority than the claims which others make about us. What is more, if one is in a position to report authoritatively on what is on one's mind, one is usually right. What is the source of the authority these statements carry? And what explanation can be given of their frequent convergence with the psychological states that others are apt to ascribe to us third-personally?

The most deflationary answer that can be given is that offered by Richard Rorty (1981), who supposes that it is a convention of our

end p.409

language that we respect what people say about their own minds. This makes the authority granted to the speaker's self- pronouncements merely a courtesy of the interpreter, not a fact about the speaker's standing with respect to his own psychology. Notice that deferring to speakers in this way ensures the truth-conditional symmetry between first- and third- person ascriptions. Granting truth to the former, we thereby settle the truth of the latter. However, this does not guarantee first-person ascriptions any real subject-matter, or explain their convergence with third-person ascriptions of the same state arrived at independently. Nor will it allow room for us to criticize someone's self-assessments on cognitive grounds. The problem of self-knowledge requires us to reconcile first-person authority with objectivity. But this is not a candidate solution, since it ensures authority by surrendering objectivity. The intersubjective agreement about these claims is simply fabricated.

But if we respect the genuine authority of subjects' views of their current mental states, and we acknowledge the fact that whenever they have such views they are usually right, we need a better account of what gives them access to their inner lives.

But what do subjects have such access to? Whatever it is, it must be the same thing that we can have access to by interpreting them. And in interpretation, 'the objects to which we relate people in order to describe their attitudes need not in any sense be *psychological* objects; objects to be grasped, known or entertained by the person whose attitudes are described' (Davidson 1987: 62). Instead, the objects to which we relate people to keep track of their thoughts are the sentences of our language, sentences which function semantically as the complements of that-clauses in propositional attitude reports. This gives us knowledge of other people's mental lives without any perceptual, or other, access to what they are thinking. But what gives the subject knowledge of his own mental life?

For Davidson, someone's being in a mental state is a matter of him being in a physical state that can be described in intentional terms. Such states are inner in the sense of being identical with particular physical states of the agent. But this will not help the person in those states to know *which* intentional states he is in. A state is individuated as the mental state it is by its intentional description. And while interpreters can know which states of mind an agent is in, because interpretation licenses particular intentional descriptions of him, this does not help the agent himself to know which intentional descriptions his states satisfy. To suppose it did, would be like his trying to read a label on his forehead without a mirror. The Davidsonian agent finds himself in this predicament, because the idea of direct access to the mind 'from

end p.410

the inside' has been abandoned. The agent has no direct cognitive contact with his inner states, no privileged access to an inner realm. What Davidson is urging on us is a radical shift in our conception of states of mind:

The central idea I wish to attack is that these are entities the mind can 'entertain', 'grasp', 'have before it', or be 'acquainted' with . . . Of course people have beliefs, wishes, doubts, and so forth; but to allow this is not to suggest that beliefs, wishes, and doubts are entities before the mind, or that being in such states requires there to be corresponding mental objects. (Davidson 1987: 61–2)

In similar Wittgensteinian spirit, Crispin Wright suggests that there is no inner epistemology of intentional states: 'knowing of one's own beliefs, desires and intentions is not really a matter of "access to"—being in cognitive touch with—a state of affairs at all' (Wright 1989a: 631–2).

Having dispensed with the inner realm, it is easy to appreciate why both these thinkers should conceive of language as what

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

relates us to our mental lives. But how does this help us to know our own minds? All that has to be explained, according to Davidson and Wright, is the first-person authority that speakers have in reporting their own mental states. First-person authority distinguishes the way we know our minds from the way others know them. But this is not some form of contact with our mental states. It is simply an asymmetry between first- and other-person psychological ascriptions of a state to a thinker, in that the former carry a presumption of correctness which the latter do not. Davidson tells us that the problem of self-knowledge can be raised 'either in the modality of language or of epistemology', and goes on to say: 'I assume therefore that if first person authority in speech can be explained, we will have done much, if not all, of what needs to be done to characterise and account for the epistemological facts' (Davidson 1984b: 102). However, the role which language plays in Davidson's strategy is stronger than these remarks suggest. We are also told that 'the only access to the fine structure and individuation of beliefs is through the sentences speakers and interpreters of speakers use to express and describe beliefs' (Davidson 1986a: 315). A thinker's self-ascriptions can be justified by the interpreting descriptions which others give of his behaviour. For Davidson, what needs to be explained is why, without the aid of these justifying descriptions, a thinker can come up with psychological self-ascriptions which are usually correct. A full account will have to address questions (1) and (2) above. So, as well as explaining their typical correctness, we will need to know what equips speakers to make these effortless self-ascriptions. A necessary part of this will be accounting for the speaker's understanding of the words used to make ascriptions.

end p.411

Wright envisages a similarly strong connection between language and self-knowledge: 'Intentional states are *avowable*: they are subject to groundless, authoritative self-ascription, and it belongs to their essence that they are so' (Wright 1989a: 629). At first sight, he seems to be committed to the claim that linguistically expressed self-knowledge is necessary even to have such intentional states. This seems too strong. But perhaps all Wright is committed to saying is that creatures with language, who can usually express their opinions linguistically, will be able to express the opinions they usually have about their states when they are in them. Creatures without language would probably lack the kind of intentional states of which we normally have authoritative knowledge. In the normal case, the avowals that subjects produce are largely correct. Normally, how things are in our minds is how we avow them to be. If this is the usual state of affairs, perhaps we need no special account of how we achieve self-knowledge. Special explanations will be necessary only when the subject is self-deceived, or for some other reason is unable to avow his beliefs, desires, or intentions. The default position is that our avowals will be true, and that our ability to avow them is what it is for us to be in those states. This is Wright's deflationary suggestion. No explanation is needed of how we close the epistemic gap between being in a mental state and knowing that we are. On this approach, the status of the judgements which subjects make about their mental states plays an important role in framing the truth about the states of mind they are in. The psychological facts of the matter are not independent of the judgements we make about them. As Wright puts it:

the authority standardly granted to a subject's own beliefs, or expressed avowals about his intentional states is a constitutive principle: something that is not a by-product of the nature of those states, and an associated epistemologically privileged relation in which the subject stands to them, but enters primitively into what a subject believes, hopes and intends. (Wright 1989a: 631–2)

The judgements which a subject makes are not always correct—so it is not the case that whatever seems right to him is right—but truth is the default assumption. These judgements can only be overturned when the interpretations we can give of someone's behaviour, which leave out or reject the subject's self-pronouncements, give better service in interpreting his linguistic and other behaviour than interpretations which respect them. For Wright, the judgements a subject makes about what he believes, desires, or intends, when made under cognitively ideal conditions, determine the truth of those claims. In this way, such judgements do not track independently constituted states of affairs; rather, they enter into the constitution of such states. In Wright's words, they are extension-determining, not extension-reflecting judgements (1989b).

end p.412

This is an attractive strategy, but it shows some strain. For instance, there is no scope on this account for the idea that what subjects say about themselves can confirm, or be confirmed by, how we had understood them up until now. Similarly, we often take what people say about themselves to ring true, given what we know about them. But people's self-ascriptions cannot be given independent confirmation by what others are prepared to ascribe to them, because the facts about which mental states subjects are in is not something settled independently of their opinions about this. There are no independent facts by which to assess a subject's opinions about himself. Which ascriptions it is correct for us to make of a subject, and which of his self-ascriptions are correct, are arrived at all at once, by providing the most internally coherent system of attitude

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

attributions that will make best sense of the subject's behaviour and try to respect the authority of his self-pronouncements. So it would be inaccurate on this picture to say that someone lacks self-knowledge because his view of himself is at odds with what we know about him. Strictly speaking, there is no sense in which someone's self-ascriptions are compatible, or incompatible, with the facts. A subject's opinions have an impact on determining the overall pattern of assignments that settle the facts about which intentional states he is in. Why he is largely right about this, on no grounds at all, is for Wright, a relatively a priori matter. It is because it is an a priori constraint on any successful interpretation of a subject that it regards truth as the default assignment for his avowals.¹⁴

¹⁴ It is interesting to note that although Wright sees first-person authority—the default truth of status of avowals—as an a priori constraint on successful interpretation of others, he also speaks of it being 'a deep contingency' that interpretations which respect a subject's self-pronouncements give better service in making sense of him than those which overturn them. But notice that the contingency here is substantial, and therefore the case is unlike the combination of contingency and apriority that Kripke envisages in the case of the standard metre bar being one metre long. Here, the contingent fact is created by convention. Self-knowledge should not be treated as this kind of contingent a priori knowledge unless one is attracted, like Rorty, to the idea that its products do depend on convention.

We must always strive to accommodate as much of the subject's view of himself as we can. What we are inclined to think about someone cannot confirm, because it includes as evidence, what they say about themselves.

Another puzzling feature of Wright's account arises because it is possible to have intentional states without having knowledge of them. As Wright acknowledges (1989a: 633 n.), our judgements regarding our own intentional states are only sufficient (by default) for us to be in those intentional states, not necessary. So in the case where we lack self-knowledge, the presence in us of intentional states owes nothing to our judgements about them. Why, then, should it be that in cases where I

end p.413

do know my own mind, my judgements help to determine, or perhaps constitute, which states I am in? Why is the presence of intentional states dependent on my judgement in one case, but not in the other?

Leaving these concerns to one side, we should note that the great virtue of this account is that it serves to explain why it is that when we think we are in a particular state we are usually right. However, I said above that any satisfactory solution to the problem of self-knowledge, as well as explaining why we are usually in the psychological states we take ourselves to be in, should explain what it is for us to take ourselves to be in any given state in the first place. Wright offers little on this, save to say that we are 'inundated day by day with opinions for which truth is the default position', and that 'we are each of us ceaselessly but—on the proposed conception—*subcognitively* moved to opinions concerning our own intentional states which will give good service to others in their attempt to understand us' (1989a: 632, emphasis added). They spring fully formed from the subject's lips with nothing to guide them but causal forces. We produce them without forethought. But even if there is no story of why subjects produce the particular judgements they do, we can still ask how they understand these self-ascriptions? If a subject finds himself merely subcognitively compelled to produce sounds from time to time, which it transpires are true by default, and thereby fix the truth of claims about his first-order states, this will not allow room for a genuinely first-person point of view.

To understand our self-opinions is to know what we mean by our avowals. This must amount to more than just mouthing something subcognitively produced. We need some account of our first-personal comprehension of what we are saying. But where in Wright's approach are the resources for this? The materials are to be found in work by Wright (1989b) in which he examined the appealing idea that meanings are secondary qualities and share the same epistemology. Here too, our opinions, under cognitively ideal conditions for judgement, determine the truth of claims about what our words mean, or where they are to be applied. However, the details of this account make ineliminable reference to the intentions of speakers, and their judgements of meaning, and this introduces a threat of circularity into the linguistic strategy's attempt to explain self-knowledge. Perhaps, it could be said that we do not know the intentions with which we use our words. But this would deprive us of the effortlessly authoritative knowledge we have of the meanings of our words. This does not close off the options for this strategy, but more is needed.

Let us turn, at last, to Davidson's approach. It is a clear instance of the linguistic strategy. The account of first-person authority is an account

end p.414

of an asymmetry between speaker and interpreter, derived from the general conditions for interpreting a person's speech

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

and actions. To know what a person believes, one must know which sentences he holds true. We can identify his attitude towards a sentence even though we do not know what the sentence means. We know that someone will take a sentence, say *S*, to be true because of what *S* means in his language, and what he believes to be the case. So, if we know what the sentence means, we know what he believes to be the case. While, if we know what he believes, we know what the sentence held true means. Instead, we have to try to solve for two unknowns simultaneously: this is the programme of radical interpretation. Davidson points out, however, that in the general scheme of interpretation, when it is the case that a speaker holds *S* true, and knows what *S* means, it is only the speaker who is in an immediate position to know what she believes. For the interpreter to find out what the speaker believes, he will have to interpret *S*. But this is unnecessary for the speaker, who already knows what she means by *S*. And because speakers know what they mean, they will have one less hurdle to go through in moving from holding sentences true to knowing what they believe. This puts them in a better position to know their own minds. In this way, Davidson explains the asymmetry of firstperson authority by means of a linguistic asymmetry between speakers and interpreters. Speakers have to be interpreted by others, but they do not have to interpret themselves. They quite literally speak their minds, and in understanding what they say, they immediately know what they think.

But how well does this explain *first-person authority*? It certainly points to an asymmetry, but the linguistic asymmetry it is based upon appears to remain unexplained. Is Davidson resting an explicable asymmetry on one that is inexplicable? Not quite. The linguistic asymmetry is based on the difference between answers given to the questions:

- (a) How do we know what others mean by their words?
- (b) How do we know what we mean by our words?

The answer to (a) is that we know by interpreting them. In answer to (b), the key point is that we know what we mean without interpretation. Speakers do not, and could not coherently, interpret themselves. But the question we must now ask parallels (I) above:

- (II) How do speakers know what they mean, given that they do not interpret themselves?

The answer to this question is important, since it will explain what puts them in a better position than their interpreters to know their own minds.

end p.415

In answer to (II), Davidson tells us that speakers could not fail to know what they mean by their words. They must be credited with knowing what their words mean because we cannot suppose that speakers are usually in error, or continually misapply their words, without thereby supposing that they are uninterpretable, and therefore unintelligible. In so far as someone is interpretable, he must be using his words correctly more often than not. This is why, to be competent speakers, we must typically mean what we do by our words.¹⁵

¹⁵ As another sign of the close affinities between Davidson's thinking and that of Wittgenstein in the later writings, consider the following passage from Wittgenstein (1992: 36e): 'The inner is hidden from us means that it is hidden from us in a sense in which it is not hidden from *him*. And it is not hidden from the owner in this sense: *he utters it* and we believe the utterance under certain conditions and there is no such thing as his making a mistake here. And this asymmetry of the game is brought out by saying that the inner is hidden from someone else.'

Does this constitute an answer to (II)? It tells us that we usually do know what we mean by our words, but not *how* we know this. It can be agreed that our words usually do mean what we take them to mean, and that they must do. But what is it for *us* to take our words to mean something in the first place? Surely, it is intelligible to ask: How do we know what our words mean, *given that we usually do*? When the corresponding question was asked about thought, it was treated as an explanatory, rather than a sceptical, question which required a substantial answer. But we see a subtle shift in Davidson's handling of the two cases. When asked how we could know our minds without interpreting ourselves, Davidson was able to say, we know our own minds because we know what we mean when we say what we think. But when we ask the corresponding question about how we know what we mean without interpreting ourselves, Davidson treats this as the sceptical question and gives a transcendental argument in response. We could not fail to know what we mean and still be speakers. But scepticism is not the issue, just as it was not the issue with respect to knowledge of our own minds. The response to the sceptic assures us that we have to know what we mean, but we can still ask: How do we know this? This is important, because what a speaker takes himself to believe is what he understands by the sentences he holds true.

Has Davidson explained how thinkers know what they are thinking without interpreting themselves? Given that there is something the speaker means when he holds a sentence true, there is something he believes. But in order to *know* what he believes, surely he has to *know* what he means. We know what this is by constructing an interpretative truth-theory for his

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

language that assigns truth-conditions to the

end p.416

sentences he holds true. But how does *he* know what he means? As an account of self-knowledge, it tells us *that* the speaker knows what he believes, but not *how* he does. There is the interpretatively guaranteed fact of meaning something, or of our words meaning something. But what kind of knowledge does this give us of what we mean and think?

For Davidson, there *is* no more to be said here. The reason has to do with his third-personal approach to meaning and mind. Apart from special circumstances, he sees no difference between someone meaning something and his knowing what he means. Of course, it is uncontroversial that these two notions should coincide. But the question is whether the former includes or deflates the latter. Davidson prefers to replace the locution 'knowing what we mean' by 'meaning what we say'. And this is part of his deflationary tactic. For we mean what we say, according to Davidson, if we can be *interpreted* as meaning this more often than not. But this, of course, is to turn from the speaker's point of view, and to concentrate on the third-personal perspective of the interpreter. Speakers, we will be told, mean what they do by their words because those words are regularly applied to the same things. But in applying them they do not act blindly. We do not just use words: we understand them. Speakers know what they mean by the sentences they hold true. This is fixed by the truth-conditions they take them to have, and it is this knowledge that is available to them without interpretation. So what does a speaker's knowledge of truth-conditions consist in? ¹⁶

¹⁶ This is something that Michael Dummett has repeatedly called for an account of. See Dummett 1975, 1976.

Once again, there seems to be a missing level of knowledge.

It is as if no more is required of the speaker than for him to produce movements and sounds which, by the interpreter's lights, turn out to be regularly interpretable as meaning one thing rather than another. But where in this picture is there room for the immediate understanding of our words? The phenomenological experience of hearing the meaning in someone's words is undeniable. We do not first hear sounds which we then have to make sense of. Normally, uptake is immediate: we just hear what someone is saying. It is also involuntary. If someone speaks audibly in a language I understand, I have no choice but to hear their words as meaning what they do. What people say can get straight through to my mind uninvited. ¹⁷

¹⁷ The phenomenon is stressed, though for quite different reasons, by both John McDowell and Jerry Fodor.

I am not suggesting that this inner experience of comprehension should be construed as a private or wholly inner act. The meaning and understanding of a speaker's language should be publicly ascertainable. But we are missing something crucial to the

end p.417

first-person perspective of the language-user if we do not recognize a sense of comprehension beyond that of the speaker producing words the *interpreter* can find meaningful. However, the immediate comprehension of speech depends on more than just a subjective impression of meaning. It amounts to the inner recognition of public and objective facts. (This makes it the perfect place to locate our knowledge of our own minds.) The task, in trying to account for the first-person point of view of the language-user, the point of view from which he knows his own mind, is to see how the phenomenology can deliver knowledge and why it is also an epistemology of understanding. ¹⁸

¹⁸ I am entirely in sympathy with Jonathan Lear's (1984) insightful remarks, which I have found helpful here, and with the way he sets up the problem of meaning and understanding in the later Wittgenstein.

We cannot just resort to unexplained notions of understanding, or first-person knowledge of meaning, to say what it is for the speaker to mean something by his words. On the other hand, we cannot settle for such a resolutely third-personal view of what it is to possess a language, where this is mainly a matter of how one is interpreted. Davidson is not unaware of the need to say more about how the speaker understands the words he uses. However, all we are told about what the speaker means by the sentence he holds true is 'that in reporting on his understanding of a sentence, the speaker "cannot improve on" on the claim, "my utterance of the sentence 'Wagner died happy' is true if and only if Wagner died happy"' (1984b: 110). But precisely what *piece of knowledge* is expressed in the report the speaker cannot improve upon? What is expressed by an instance of the disquotational schema, and how substantial, or otherwise, is this knowledge?

Davidson will insist that there *is* more to the speaker's meaning something than his just uttering words that others regard as meaningful. In particular, he will point out that speech is an intentional activity, and that the speaker must intend to be understood in a particular way. There is a worry, here, however, if one has to know one's intentions in order to understand the

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

words used to speak one's mind; after all, thinkers usually do know what they intend, as part of having first-person authority. Likewise, if one has to know how one intends to be interpreted, one may have to know what could be discovered by an interpreter of one's behaviour, and this amounts to knowing what an interpreter would be justified in ascribing to us on the basis of our behaviour. But how, without appeal to a theory, does the speaker know this? This threatens to make knowing what we mean a highly reflective matter of knowing how we could be interpreted. However, Davidson avoids these worries.

end p.418

We are told that to have reasons, or to act on intentions, 'the speaker does not usually "form an express intention" and he does not "hold a theory"'. Further, 'intentions are not normally attended by any special feelings' (1992: 259). It is enough for the speaker to be credited with the intention of being interpreted in a certain way, and it is enough for someone to be acting 'intentionally when there is an answer to the question what his reasons in acting were', where this is a matter of the intelligible sense which the interpreter can make of his reasons. Once again, Davidson is doing a highly laudable job of trying to accommodate the first-person perspective, but he shifts in his description of that predicament to what the interpreter would ascribe or recognize. What is not being acknowledged sufficiently here is that reasons have insides as well as outsides. They are reasons the subject can appreciate from his own point of view. To suppose otherwise is to leave the Davidsonian agent in the position of Hamlet as seen by Goethe's *Wilhelm Meister*, who thinks of the prince as saying: 'I am in a play full of purpose but have none of my own'.

If the Davidsonian strategy cannot accommodate the first-person perspective of the language-understander, it will ultimately fail to make room for the first-person point of view. As a strategy, it offers us genuine insight into the connections between knowing one's mind and knowing one's language; but more is required to explain what it is to know one's language. In so far as an account can be given which preserves the first-person point of view of the language-user, and first-person awareness of meaning, we shall be able to use the strategy illuminatingly; but in so far as these features of the language-user's predicament are missing, or neglected, the account will lack the materials to explain the first-person perspective on our minds.

To recap: the linguistic strategy is the idea that we know our states of mind by making largely correct and authoritative claims about them in a language we understand. We need an account of what equips us to produce and understand the claims we make about what we are currently thinking, and an explanation of why what we think we are currently thinking is usually a good guide to what we are thinking. In their different ways, Davidson and Wright give answers to the second question. But an account of self-knowledge is not complete until we have a satisfactory answer to the first question.

Where are we to find the resources to account for the first-person point of view of the language-user, the point of view from which he knows his own mind? I have suggested that more is needed to explain how speakers know what they mean. This involves the immediate experience of hearing, or understanding, utterances as meaningful. But how

end p.419

are we going to cast light on this immediate first-person awareness of meaning? We are faced with a dilemma. If we draw upon first-person aspects of our psychology, the sort of states for which we have first-person authority, this threatens to make knowledge of what we mean depend on knowledge of our own intentional states. So the strategy of explaining knowledge of one's own mind via knowledge of one's language will be circular. On the other hand, if we prescind from any notions in the domain of personal-level mental states, we will lack the resources to explain first-person aspects of linguistic understanding, and hence knowledge of what we are thinking.

On the second horn of the dilemma, we may suppose that some highly modular part of the brain is dedicated to handling language. Call this the language faculty. It is this faculty which is called upon to handle speech produced in the expression of judgements we make about our current states of mind. The modular view of the language faculty avoids circularity in the linguistic strategy because it amounts to one part of the mind—a subpersonal part—being put to service in understanding another. However, it is hard to see why what is competently handled by the underlying linguistic system can deliver instances of self-knowledge to the subject, and contribute to elaborating the notion of a first-person point of view. Alternatively, on the first horn, if we reach for more states of the subject at the personal level to flesh out a conception of the speaker's first-person awareness of meaning, we risk helping ourselves to notions we are trying to explain.

So what is the nature of our knowledge of meaning? Ordinary reflection gives us few clues, except to confirm that we are effortlessly authoritative about what our words mean, whether word strings are grammatical sentences, and how words in our language are pronounced (Higginbotham 1991). What does this *first-person linguistic authority* consist in? For instance, what

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

is the knowledge reported in the disquotational scheme? Is disquotational knowledge the best we can do to cast light on what a speaker knows?

In a way, disquotation is trivial. In using expressions to state their own meanings, we cannot fail to express a piece of linguistic knowledge. But this is only the case so long as the expressions as disquoted mean something to us. Genuine instances of disquotation merely reflect the familiar fact that, as speakers, we cannot be alienated from our own language and the meanings of our words. This is relatively trivial. But what it takes for something to be a genuine instance of the disquotation scheme is strictly non-trivial. So I agree with Akeel Bilgrami, that 'Disquotation, if it is to be at the service of an account of meaning is not a wholly trivial idea. It must be anchored in something which is not made explicit in the disquotational clause itself' (1992: 247).

end p.420

What is required for something to be a genuine instance of the disquotational scheme for a speaker? Obviously, we need to understand the sentences featured as they are used. And since speakers do not carry knowledge of sentences around from one occasion to another—sentences, like thoughts, are ephemeral—a speaker's knowledge of meaning and grammar must be at work in delivering an understanding of a sentence, quoted on the left-hand side and used on the right. If we did not exercise knowledge of grammar, we could not tell which assemblies of words were sentences, or what compositional structure the sentences had. The standing linguistic knowledge we carry from one occasion to another consists in the words and means by which we produce and comprehend anything that counts for us as a meaningful utterance.

In accordance with Chomsky's notion of an I-language (Chomsky 1986), I am taking a grammar to be an internalized system of representations, or body of knowledge, whose workings are not known to the ordinary language-user, but which have an effect on his conscious recognition and production of speech. My grammar will assign structure to strings of signs or sounds which will determine the sentences or expressions I hear. I will typically hear strings as structured, and thus be able to recognize them as utterances of particular sentences. The sentences I hear will depend for their identity on the internal states I am in, the structural assignments that the language faculty makes to sounds or signs I encounter. This is because, as Higginbotham puts it, 'the crucial properties of sentences are not revealed by thinking of them as they are outwardly presented to us, namely as strings of signs, but rather by their unobservable grammatical structure' (1991: 555). Which sounds or signs are sentences for me is a matter settled entirely by my internal linguistic system. The same sounds or signs could be interpreted quite differently by others, or not count as sentences at all. 'What it is for something to be a sentence for a person is for it to be a grammatical structure that is apprehended and applied to certain perceptible objects' (ibid. 556). Only if the speaker has the linguistic competence to confer lexical meaning and sentence structure on the string of words quoted in any purported instance of the disquotational schema will this produce a true instance of the schema. Not everyone agrees that even *this much* is needed.

Higginbotham (1989), for instance, thinks the following counts as a genuine instance:

(T1) 'All mimsy were the borogoves' is true-in-L iff all mimsy were the borogoves.

Would this count as knowing 'the first thing' about the meaning of the sentence featured in (T1)? Higginbotham says we would know its logical form, in skeletal terms, and that the categories of the words that

end p.421

occur in it typically play certain semantic roles in the language. When we know how the parts of the sentence are combined to form something with that skeletal form, we know its compositional structure up to the meanings of certain lexical items. What is lacking is just knowledge of the lexical axioms. But these can be provided as follows:

(L1) 'mimsy' means mimsy (*whatever that is*);

(L2) 'borogoves' means (or refers to) borogoves (*whatever they are*).

Higginbotham concludes that this is all one knows of meaning from disquotational knowledge, and he is happy to accept that disquotational knowledge of truth and reference is trivial. We could, as speakers, go in for explications of word meanings, but these may be quite unreliable guides to use, and do not (or do not wholly?) fix meaning, or reference.

By my lights, not only am I unable to know which (as opposed to which kind of) proposition is expressed by (T1), it fails to express a truth about my language, and therefore is not, as I utter it, a genuine instance of the disquotational schema. For me, whether any featured word string gives rise to a genuine instance will be a matter of whether my language faculty, together with my knowledge of word meaning, can assign the string of words semantic, syntactic, and phonetic properties that determine it as a meaningful sentence of my language, or idiolect. It will be a fact about me, whether I understand a sentence and can report my understanding in this disquotational form. Notice that the disquotational form may be a genuine

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

instance for one speaker but not for another; and for the same speaker, two different pieces of disquotational knowledge can be reported by one and the same form when the featured string is ambiguous.

For Davidson, on the other hand, what would establish whether there is a truth of my language expressed by any putative instance of disquotation is success in interpretation. A fully formed interpreter can derive, by means of an interpretative truth-theory for my language, a T-theorem corresponding to the form of the disquotational sentence:

'Wagner died happy' is true in L iff Wagner died happy

that gives the truth-conditions of my sentence mentioned on the left-hand side. It will not necessarily be homophonic, unless the interpreter's use of those words, and my use, coincides. The theorist will give his rendering of the same piece of knowledge, available to me without the help of a theory. And on Davidson's reading of the publicity constraint, all I can mean by this sentence is all I can be known to mean as determined by the pronouncements of an interpretative truth-theory: ' . . . what the interpreter needs to know of the semantics of the speaker's language

end p.422

. . . is conveyed by the T-sentences entailed by a theory of truth' (1990: 319). But just what is expressed by a T-sentence? The piece of knowledge it records will express what I believe in holding true the sentence it mentions. Remember, an interpretative truth-theory is confirmed by the totality of evidence concerning the speaker's behaviour, linguistic and otherwise, and the only empirically significant notions of meaning and structure are those which supervene on that behaviour. Although it is not reducible to behaviour, 'meaning is entirely determined by observable behaviour. . . . That meanings are decipherable is not a matter of luck; public availability is a constitutive aspect of language' (1990: 314). And for Davidson:

what is open to observation is the use of sentences in context, and truth is the semantic concept we understand best. Reference and related semantic notions, like satisfaction, are, by comparison, theoretical concepts (as are the notions of singular term, predicate, sentential connective and the rest). There is no question of the correctness of these theoretical concepts beyond the question whether they yield a satisfactory account of the use of sentences. (1990: 300)

I suggest words, meanings of words, reference and satisfaction are posits we need to implement a theory of truth.

They serve this purpose without needing independent confirmation or empirical basis. (1984a: 222)

Thus, the content of T-sentences will not involve the theoretical notions used, purely instrumentally, within the theory to derive T-sentences. So the content of the speaker's knowledge: what he means by the disquotational report, which can also express what he believes, will be very lean indeed. What Higginbotham, at any rate, helps us to see is that there is knowledge of grammar at work in understanding the quoted sentence. Facts about grammar that determine some aspects of meaning have to be called upon to explain what goes into my understanding. This will call upon, and provide, more than the interpreter sees from the evidence he is typically exposed to. So we can give some of the items on Davidson's list empirical confirmation in terms of the properties of the speaker's grammar. These will include meaning-affecting properties of scope and the referential dependence of nominal items within the sentence, all of which constrain my semantic interpretation of sentences.¹⁹

¹⁹ For example, for reasons to do with purely syntactic binding principles, we cannot read the sentence 'John shaved him' as referring to the same person in both positions, whereas 'John's father shaved him' may be understood as John's being shaved by his father, but not his father shaving himself.

But there is more to understanding than this. As Michael Dummett stresses, language is a conscious and rational activity, and speakers are consciously aware of what is said. Certainly, language understanding

end p.423

depends on the unconscious workings of my internal language system. But there is a product delivered to consciousness. The purpose of conscious experience, from a cognitive processing point of view, is that it functions to present a single interpretation of a diversity of information. For example, we cannot see both faces of a Necker's cube at once; nor do we hear two (or more) readings of an ambiguous sentence string simultaneously. We always hear one reading, and recognize that another reading is available. But we are never presented with something that encompasses both or is neutral between the two. We always hear the string as an utterance of one particular sentence rather than another, even though the string of sounds or signs may be assigned more than one reading—may be seen or heard as the utterance of one or another sentence type.²⁰

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

²⁰ I doubt whether sentences have tokens, although utterances and inscriptions do. These will be typed by sentences and expressions.

These events of perceptual awareness in speech processing depend counterfactually on the underlying states of the language faculty and which structural assignments were made to which strings. Had my grammar assigned another structural description to this string, I would have heard another sentence uttered. We know this because we know how recent exposure to other structures affects the internal states we can be in when hearing the next sentence. For example, reading the sentence 'Time flies like an arrow' immediately after reading the sentence 'Fruit flies like a banana' will make one immediately aware of another construal of the first sentence, and a previously unnoticed ambiguity. This depends on our linguistic system assigning different grammatical categories and sentence structure to the string we read.

What are we to make of all this in trying to explain a speaker's first-person linguistic authority, and how can we avoid the dilemma faced by the linguistic strategy?

For the case of our knowledge of grammar we simply steer between the horns of our dilemma. First-person authority about which sentences are grammatical can be secured by adopting what Higginbotham (1991) calls *representationalism*. This is the view that in language, how you represent things is how they are. How things are, grammatically, will be a matter settled by the internal states of my language faculty. Whether or not a string of words is a sentence for me, will depend on whether I, or, rather, my language faculty, can assign it a structural description that makes it well formed. If it can, I shall recognize it as a sentence. If not, I shall simply hear it as deviant. Which strings of signs or sounds count as normal or deviant will differ from speaker to speaker. There is a huge variation across speakers of what is called English. For

end p.424

me, the pair, 'I am/Amn't I' is normal, whereas the (curiously?) plural form in the pair, 'I am/Aren't I' sounds odd. Many of you will have the opposite reaction. There is no standard of right and wrong here, save what is produced in accordance with our own internalized grammars. And as ambiguous strings show, the grammatical structure of sentences is not in the sounds or signs themselves, but in our perception of them. I do not have conscious knowledge of the internal facts in my linguistic system, but I am aware of the effects of this tacit knowledge on me as a conscious language-user. In language, how you represent things grammatically is how they are. Hence we are authoritative about our own grammar. However, this is not to surrender the objectivity of grammatical judgement. We need not lose sight of the distinction between how things seem and how they are, because I do not have perfect knowledge of how my internalized linguistic system represents things. I am not entirely blind to its properties, because they make a conscious impact on me delivering one rather than another parsing of a syntactic string. But other factors, such as memory and attention, influence my speech processing. So I am not infallible. What is more, linguists can construct and test hypotheses about the grammar of my idiolect. So there is an independent perspective on the facts.

But if we all speak different idiolects, how can we understand one another? The crucial consideration here is whether speakers have largely shared, or overlapping, vocabularies. Their grammars will differ, but this will be less important. As Chomsky (1986) has shown, grammars vary within strict parametric limits, since speakers are innately endowed with a Universal Grammar whose principles are common to all human languages. Universal Grammar describes the initial state of the language faculty of every speaker. Its principles are parameterized, and set locally by the environmental influence which other speakers have on the child. The people we are most in contact with, and from whom we learn our vocabulary, will have grammars that overlap with our own. And even where there is greater variation, of the sort we often encounter when communicating with non-native speakers of a language, influenced by the parameters of their mother tongue, we usually have enough to go on to know what they mean when their grammatical constructions use words that are familiar to us. Thus, I do not need to speak as they do to know what they mean, although the closer people's languages are to mine grammatically, the easier and more direct my understanding of them.

What is important is that we have largely shared knowledge of the meanings of words. However, knowledge of word meaning is not part of the underlying linguistic system, and is not, like representations of

end p.425

structure, entirely a matter of what is internal to me. Knowledge of word meaning is knowledge of the norms that govern our use of words. These are public norms which introduce third-personal standards to our first-personal use. It is also the case that the meanings of many of our words are fixed, as externalism suggests, by reference to entities in our environment. Yet, we have immediate authoritative knowledge of what words mean. This raises the familiar issue of how to reconcile first-person authority with objectivity. And here, unlike the case of grammar, we still await a satisfying account of how this is to be done.

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

Knowledge of word meaning belongs within the conscious rational mind. Command of word meaning, as McDowell insists, is a conscious mental ability. So these are first-personal dimensions to our understanding which remain unexplained. This leaves a lacuna in the linguistic strategy, and it is only when we have an explanation of the phenomenon of first-personal, authoritative knowledge of meaning that does not draw upon non-linguistic elements of our knowledge of our own minds that we can provide a non-circular explanation of our psychological self-knowledge. So there is work still to be done.

The first part of this task is to say how we can be effortlessly authoritative about the objective standards governing the proper use of words. This is, of course, just the issue Wittgenstein faced in the rule-following considerations, and to which he did not have a fully satisfying answer. Most of the elements of the problem of self-knowledge occur here. We use words without justification, and yet, as Wittgenstein tells us, 'To use a word without justification is not to use it without right' (1953: §289). What counts as using a word correctly is settled, for him, by shared standards of communal use. But this leaves him to ask: how can the meaning of a word seem to be all there at an instant when it is also the unfolding of use through time? (see §138), a question to which he gives no satisfactory answer.

In answering this question, McDowell would suggest that the meanings are there to be perceived on the surface of our communal practices, described in content-involving terms. These facts are available only to those who have acquired membership of this shared linguistic practice. But there is no good empirical evidence that speakers do share a common language. Nor do they need to in order to communicate successfully. Furthermore, McDowell's epistemology of language understanding mislocates elements of this understanding, such as our knowledge of syntax, on the surface of our practices where it simply cannot belong. There may be no unified linguistic object, no single locus of linguistic significance. To appreciate word meaning is to be responsive to something out there, but responses to our own assignments of structure may

end p.426

dictate much of what we hear said. And this points to another role for the conscious, first-personal perspective of the speaker to play. The experience of conscious comprehension is required to unify, in the first-personal realm, elements from the third and subpersonal domains which all have their impact on linguistic understanding. The form and character of the conscious products of understanding are shaped by these factors. When we ask, finally, what do we know when we know a language, the conscious experience of understanding tells us that it will not be the indeterminate or inscrutable product that the interpreter assigns to us. What is given in genuine instances of disquotational knowledge must be much richer than the interpreter's T-theorem assignments suggest, and only when we know more of this, can we call upon what we know when we know our own language to explain how we know our own minds.

References

Bilgrami, A. (1992), 'Can Externalism be Reconciled with Self-Knowledge?', *Philosophical Topics*, 20: 233–67.

Burge, T. (1988), 'Individualism and Self-Knowledge', *Journal of Philosophy* 85: 649–63; repr. in Q. Cassam (ed.), *Self-Knowledge* (Oxford: Oxford University Press, 1994), 65–79.


Chomsky, N. (1986), *Knowledge of Language* (New York: Praeger).

Davidson, D. (1984a), *Inquiries into Truth and Interpretation* (Oxford: Oxford University Press).


— (1984b), 'First-Person Authority', *Dialectica*, 38: 100–11.

— (1986a), 'A Coherence Theory of Truth and Knowledge', in E. Lepore (ed.), *Truth and Interpretation* (Oxford: Blackwell), 307–19.

— (1986b), 'A Nice Derangement of Epitaphs', in E. Lepore (ed.), *Truth and Interpretation* (Oxford: Blackwell), 433–46.

— (1987), 'Knowing One's Own Mind', *Proceedings and Addresses of the American Philosophical Association*, 60: 441–58;  repr. in Q. Cassam (ed.), *Self-Knowledge* (Oxford: Oxford University Press, 1994), 43–64.

— (1989), 'The Myth of the Subjective', in M. Krausz (ed.), *Relativism: Interpretation and Confrontation* (Notre Dame, Ind.: Notre Dame University Press), 159–72.

— (1990), 'The Structure and Content of Truth', *Journal of Philosophy*, 87: 279–328. 

— (1992), 'The Second Person', *Midwest Studies in Philosophy*, 17: 255–67. 

Knowing Our Own Minds

Wright, Crispin (Editor), Fellow of the British Academy and Professor of Logic and Metaphysics, University of St Andrews

Smith, Barry C. (Editor), Lecturer in Philosophy, Birkbeck College, London

Macdonald, Cynthia (Editor), Professor of Philosophy, University of Canterbury

Print publication date: 2000, Published to Oxford Scholarship Online: November 2003

Print ISBN-13: 978-0-19-924140-8, doi:10.1093/0199241406.001.0001

Dummett, M. (1975), 'What is a Theory of Meaning?', in S. Guttenplan (ed.), *Mind and Language* (Oxford: Oxford University Press), 97–138; repr. in Dummett, *The Seas of Language* (Oxford: Clarendon Press, 1993), 1–33.

— (1976b), 'What is a Theory of Meaning? II', in G. Evans and J. McDowell (eds.), *Truth and Meaning*, (Oxford: Oxford University Press); repr. in his *The Seas of Language* (Oxford: Clarendon Press, 1993), 34–93.

end p.427

Higginbotham, J. (1989), 'Knowledge of Reference', in A. George (ed.), *Reflections on Chomsky* (Oxford: Blackwell), 153–74.

— (1991), 'Remarks on the Metaphysics of Linguistics', *Linguistics and Philosophy*, 14: 555–66. [Link](#)

Lear, J. (1984), 'The Disappearing "We" ', *Proceedings of the Aristotelian Society*, suppl. vol. 58, 219–42.

McDowell, J. (1984), 'Wittgenstein on Following a Rule', *Synthese*, 58: 325–63. [Link](#)

— (1987), 'In Defence of Modesty', in B. Taylor (ed.), *Michael Dummett: Contributions to Philosophy* (Dordrecht: Martinus Nijhoff Publishers), 59–80.

— (1991), 'Intentionality and Interiority in Wittgenstein', in K. Puhl (ed.), *Meaning Scepticism* (Berlin: de Gruyter), 148–69.

Moran, R. (1988), 'Making Up Your Mind: Self-Interpretation and Self-Constitution', *Ratio*, 1: 135–51.

Rorty, R. (1981), *Philosophy and the Mirror of Nature* (Oxford: Blackwell).

Wittgenstein, L. (1953), *Philosophical Investigations*, trans. G. E. M. Anscombe (Oxford: Blackwell).

— (1992), *Last Writings on the Philosophy of Psychology*, ii (Oxford: Blackwell).

Wright, C. (1989a), 'Wittgenstein's Later Philosophy of Mind: Sensation, Privacy and Intention', *Journal of Philosophy*, 78: 622–34; repr. in K. Puhl (ed.), *Meaning Scepticism* (Berlin: de Gruyter), 126–47.

— (1989b), 'Wittgenstein and the Central Project in Theoretical Linguistics', in A. George (ed.), *Reflections on Chomsky* (Oxford: Blackwell), 233–64.

end p.428
