

# The Authority of Avowals and the Concept of Belief

Andy Hamilton

## 1. The Dispositional Model of Belief and the Authority of Avowals

The concept of belief is profoundly puzzling. It has affinities both with conscious states and with dispositions, and as a result some writers have wanted to analyse it into two distinct concepts. The conception of belief as judgment, feeling or mental act – as a mode of conscious thought – persisted at least from Descartes and Hume to Russell and Ramsay.<sup>1</sup> Ryle and Wittgenstein effectively undermined this conception, opening the way for the present consensus that belief is a complex behavioural disposition, or a functional or informational state underlying such a disposition – or, most usually, a combination of these. I will refer to the consensus as the *dispositional model*.

Many of those who subscribe to the dispositional model want to sever the connection between belief and consciousness entirely. Dennett for instance seems to hold that beliefs, in contrast to opinions, are merely patterns of behaviour: ‘My opinions can be relied on to predict my behaviour only to the degree, normally large, that my opinions and beliefs are in rational correspondence’. In self-deception ‘one behaves one way while *judging* another. One’s behaviour is consonant with one’s beliefs “automatically”’.<sup>2</sup> From the dispositional model arises a characteristic epistemology: knowledge of one’s present beliefs involves a generally reliable self-monitoring or self-scanning.<sup>3</sup>

The present article is concerned to reject this picture. The dispositional model is unable to recognise some essential aspects of the concept of belief. Beliefs express a commitment to truth, and involve being right or wrong on some matter; appeal to a behavioural disposition goes only a little way in acknowledging this normative aspect. The dispositional model’s essentially third-personal stance fails to do justice to the concept of a unified rational subject, thus encouraging a kind of self-dissociation. My principal focus, however, is on failings concerning the epistemology of belief. The model does not allow that the subject is fully authoritative concerning their own present beliefs. In avowing a belief, what I monitor, if anything, is the world, not a state of myself.

The expression ‘first-person authority’ is unfortunate, suggesting a kind of expertise or privileged access. I will refer instead to *the authority of avowals*, since it is preferable to regard the avowal – the first-person utterance – as authoritative, rather than the subject. The phenomenon is captured by the *authority principle*:

If X truthfully, attentively, comprehendingly, and non-self-deceptively asserts or avows 'I believe that  $p$ ', then the truth of 'X believes that  $p$ ' is guaranteed (where the sense of 'X' and ' $p$ ' is constant, appropriate indexical changes being made).<sup>4</sup>

The principle expresses the idea that brute error in avowals of belief is not possible; any error in the avowal implies some defect in the subject. By saying that the avowal is authoritative, I mean simply that it settles the question of what the subject believes in the way defined by the authority principle.

The essential thought underlying the principle can prove elusive. Here is a simple statement of it. Imagine a situation where the subject's avowal appears inconsistent with other linguistic and non-linguistic behaviour. The authority principle rules out one possible explanation of this inconsistency: that the subject is simply ignorant of the belief they hold. This possibility could not be ruled out in the case of ordinary behavioural dispositions to drive over the speed limit or smoke cigarettes, and character- and personality-traits such as generosity and obstinacy. In such cases the subject may, non-pathologically, lack insight; there is no authority, and no avowal in the ordinary sense.

The principle refers to avowals or assertions, not simply to utterances of the words – otherwise additional qualifications would be required to rule out quotation, utterances in plays, embedded uses of 'I believe that  $p$ ', for instance in conditionals, and so on. 'Truthfully' really is the condition intended here, though 'sincerely' would serve; the principle is an empty tautology only if all truthful utterances express a truth, which clearly they do not.

There is an associated principle, also in conflict with the dispositional model, which must be acknowledged in an analysis of belief. It may be termed the *decision principle*:

Having decided the question whether  $p$ , there is, for me, no further question whether I believe that  $p$ .

That is, if someone asks me whether I believe that  $p$ , I will attend to whatever I think is required to decide the question whether  $p$ . There is nothing else to which I could reasonably attend; in particular, I do not attend to my own inner states or behaviour. Thus it would be absurd for me to say, on being asked whether I believe there will be a military coup in Russia, 'Here's my provisional answer, but I can't tell you for sure until the verdict is in on my behaviour'. As Gareth Evans put it, defending a version of the principle, 'in making a self-ascription of belief, one's eyes are, so to speak, or occasionally literally, directed outward – upon the world'.<sup>5</sup> In fact, as Evans does not say, having decided whether  $p$ , there is no further question for me whether I believe that  $p$ . ('Self-ascription' tends to suggest that there is, which is one reason why 'avowal' is preferable.)

Alternatives to the dispositional model – notably that implicit in Crispin Wright's constitutive account of self-knowledge – have recently received some

attention. But despite this, in its application to belief the authority principle is still not widely held.<sup>6</sup> It is inconsistent with physicalism, often described as the only possible position in the Philosophy of Mind (though it might be truer to say that 'Philosophy of Mind' describes that branch of contemporary philosophy concerned with physicalist solutions to an alleged problem of the relation between mental and physical). But there are two other popular standpoints from which the authority principle is claimed to be untenable.

The first is *externalism*, the view that psychological content is in important respects socially and environmentally determined. Concerning these respects, it seems that the subject's avowals cannot be authoritative.<sup>7</sup> As traditionally presented, the externalist model perpetuates a crude dichotomy between inner and outer – the definition of 'internal' as 'in the head' neglects any conceptual connection between mental states and behaviour. (Its prevalence suggests that contemporary Philosophy of Mind should be re-branded as the Philosophy of Heads.)<sup>8</sup> But when the dispositional model and its associated epistemology are rejected, the threat to authority posed by externalism is dispelled – or so I will later argue.

A second popular standpoint is that of Davidsonian rational interpretation. *Holism of the mental* seems to conflict with the authority of avowals. If ascriptions of belief are not made singly but in belief-desire pairs revisable subject to constraints of rationality – for instance if the obvious consequences of *p* turn out to be implausible or inconsistent with other beliefs attributed – then it seems that sincere avowals may be overridden. Perhaps also, instead of self-ascribing present beliefs singly and directly – that is, avowing them – I aim at an overall rational interpretation of my own behaviour. I will argue that this is a mistaken picture; but holistic considerations point to a more basic reason for rejecting the authority principle.

This is what I will term *the dispositional dilemma* – though it arises if belief is construed as any kind of continuous state beyond the subject's immediate knowledge.<sup>9</sup> It is hard to understand how I could have what seems to be immediate and certain – or even, in Davidson's terms, presumed – knowledge of a state extended in time, a state which others apparently come to know of by interpreting my linguistic and non-linguistic behaviour. As Davidson puts it, the possibility of error by the subject follows, despite the presumption of correctness, from the fact that attitudes are dispositions that manifest themselves in various ways, and over a span of time. Hence reliabilism: for Davidson and many other writers too, it seems that even sincere, attentive, non-self-deceptive avowals can be mistaken – unless self-deception is to become quite implausibly extensive.<sup>10</sup>

So reliabilism seems forced upon us. For surely, whatever else it is, a belief must be a state extended in time, implying a disposition to behave? If the authority principle denies this, it must be unacceptable. The remainder of this article will be concerned with the defence of the authority and decision principles, and with the resolution of the dispositional dilemma which arises out of that defence. I will argue that avowals of belief are expressive and not descriptive. The dispute with Dennett is not merely semantic, so that what he calls

judgment or opinion, I call belief. Judgment or opinion, and belief, are inextricably linked.

## 2. 'Unconscious Beliefs': Arguments Against the Decision Principle and the Authority of Avowals

The dispositional model implies that there are *unconscious beliefs*, the existence of which undermines the authority and decision principles. For instance, Peacocke argues that if we allow such beliefs, we

'cannot then accept that if an application of my procedure for answering the question of whether  $p$  generates a negative answer or no answer at all, then I do not believe that  $p$ . A negative answer and no answer at all are equally consistent with my having an unconscious belief that  $p$ '.<sup>11</sup>

Peacocke is not exactly expansive on this question, and the only example he gives is obscure. But I think that as far as present-tense ascriptions are concerned, and with very limited qualifications for self-deception, his view is quite mistaken. It is misleading to divide beliefs into two categories, conscious and unconscious. *All* beliefs are liable at some time or other to be considered, and at others – that is, most of the time – to be out of ones mind.

The assumption that there are unconscious beliefs is very widely held, however, and it is necessary to consider the various sources from which it derives. First it is necessary to point out the common confusion of unconscious beliefs with beliefs about myself which are falsified by behaviour of mine of which I appear unaware. This behaviour falsifies the embedded proposition but does not show that the avowal as such is false. For instance, a South African judge under apartheid might have avowed 'I believe I am fair, impartial and non-racist in my judgments'. If he later became persuaded that, at the time of the avowal, he really was racist, it would not follow that his avowal of belief was mistaken – unless there is a more complex story involving self-deception.

Self-deception offers the first of four plausible senses of 'unconscious belief' which need to be considered:

(1) *Repressed beliefs: self-deception or unconscious motivation.* Here 'unconscious' is not an innocent description. Self-deception is the principal cognitive defect of which an avowal of belief has to be free in order to be authoritative. ('Cognitive defect' is not totally felicitous, but here implies a failure of reasoning for which the subject may be blamed.) This qualification does not import circularity into the authority principle, because the self-deceiver does not simply make an erroneous avowal of belief. Their 'mistakes' are motivated, and never clear-cut. Self-deception is not merely self-ignorance, but involves motivation by a wish, desire or fear, such that the subject is unable consciously

to acknowledge what they know or suspect, whilst in some way acting as if they did know or suspect it. Repression or avoidance of evidence for the state of affairs is implied.

Those who discuss the authority of avowals of belief rarely offer any account of self-deception, and simply assume that it rules out anything stronger than a reliabilist interpretation of authority. Many implicitly subscribe to the 'other-deception' model of self-deception, espoused by writers from Kant to Davidson, postulating the existence of an unconscious belief and hence a (motivated) error in avowal.<sup>12</sup> But although the implications of the phenomenon are hard to pin down, it does not involve a mistake about, or even ignorance of, an unconscious belief. Consequently, given that on the account offered here the essence of the authority of avowals concerns the impossibility of this kind of error, self-deception is not in tension with it.

One way of acknowledging these claims, and avoiding the other-deception model, is found in 'pretence' accounts of self-deception, which suggest that because the subject in some sense knows or suspects that *p*, their avowal could be regarded as insincere.<sup>13</sup> However, the benefit of easily accommodating the authority of avowals is outweighed by the risk of interpreting self-deception as *mere* pretence. It is therefore preferable to avoid the other-deception model by adopting a *deflationary account of self-deception*, implicit in the following presentation of its salient features (the belief that *p* may be, unoriginally, 'X believes their partner is having an affair'):

- (i) (a) Much of *S*'s behaviour, with the notable exception of their avowal, may be explained by imputing the belief that *p*; perhaps also (b) *S* has evidence for *p* (we say, for instance, '*S* must realise that *p*'; they avoid evidence for *p*, and/or rationalise such evidence)
- (ii) *S* non-hypocritically, and without intending to deceive, avows that not-*p*, and some of their behaviour may be explained by postulating this belief
- (iii) a desire or wish motivates *S*'s avowal
- (iv) *S* is in principle able to acknowledge their self-deception at some later time.

There may possibly be cases of self-deception where no explicit avowal is uttered, but clearly these do not bear on discussion of the authority of avowals. A more serious problem concerns the condition that the subject is avowing non-hypocritically and without intending to deceive. Surely this implies that they believe that not-*p*?

The deflationary account precisely avoids insisting on a fact of the matter concerning whether *S* believes that *p*, or whether they believe that not-*p*. The test for non-hypocritical or deceptive avowal is found in the distinctive later acknowledgment of self-deception. The subject does not admit simply to a mistake about their beliefs. Always their behaviour remains to some extent enigmatic to them – it was, after all, irrational.<sup>14</sup> Only if belief is construed as an informational state or disposition of which the avowal expresses merely reliable knowledge, would it seem plausible to say that *S* does clearly believe that *p*. People say: 'I guess I knew all along that my partner was having an affair. I just couldn't bring myself to

admit it'. They do not say 'I guess I believed all along, but couldn't bring myself to admit it'. This is not just because it is usually infelicitous to say 'I believed that *p*' when I can make the stronger claim that I knew it. It would be curious, more importantly, for the reason that belief implies the kind of commitment which involves being able to assert what is believed; that, perhaps, is why Freud never referred to repressed beliefs.

Thus there seems to be a divergence between knowledge and belief, with the latter having a closer tie to avowal. (This may suggest a qualification of accounts of knowledge as justified true belief.) So it is not the case that *X* clearly believes that *p* – the 'unconscious belief' – nor does *X* clearly believe that not-*p* – the 'avowed belief'. The distinct third-personal criteria for ascription of belief, normally congruent, yield conflicting results in cases of self-deception. There is no further fact of the matter concerning whether *X* believes that *p* or believes that not-*p*.<sup>15</sup>

Self-deception thus does not imply the existence of unconscious beliefs. Could it still mean that the decision principle is incorrect? Not in the way Peacocke suggests, viz.: 'A negative answer and no answer at all [to the question whether *p*] are equally consistent with my having an unconscious belief that *p*'. I may look at my past behaviour on being challenged about a sincere past-tense self-ascription of belief; or indeed a sincere present-tense avowal made in the past. But the case of self-deception does not make it intelligible to suppose that I make a present-tense avowal of belief on the basis of my behaviour, in the way that I may do for instance with emotions. Self-deception does not imply that belief is a disposition or informational state, one that is at best reliably self-monitored.

(2) *Belief that p, where 'Is it the case that p?' is a question I have never considered.* With self-deception, the beliefs at issue were unconscious in the Freudian sense of 'repressed'. Other alleged cases are innocent, though not in their philosophical implications. Perhaps it is a kind of unconscious belief when someone believes that *p* without ever having considered the question. 'Australia is not within non-stop swimming distance of the UK' and 'Dizzy Gillespie never sang at the Metropolitan Opera' were, before I began this article, such cases for me. If someone had asked 'Do you believe that Australia is within swimming distance?', I would have replied spontaneously and negatively. If they wanted to know whether I believed it before they asked, my answer would have been 'Of course I knew it, I'd just never thought about it'. Presumably I had thought about how far away Australia is, and how far people are able to swim. So maybe – perhaps unconsciously – I inferred 'Australia is not within swimming distance' from these other beliefs. The present heading therefore includes the question whether belief is closed under entailment – whether if I believe that *p*, and *p* entails *q*, then it follows that I believe that *q*.

Now it would be too strong to insist that I cannot be said to believe some proposition which has never occurred to me.<sup>16</sup> My view is that there may be no fact of the matter concerning such cases. There may be no determinate account of all that I believe, any more than there is a determinate account of all that I can see. But often the attribution of belief depends on what interpreters know or believe

– in particular, if the interpreter knows or believes otherwise. One would not attribute to someone the belief that, when he sits down, poisoned darts are unlikely to shoot up and embed themselves in his buttocks, causing a lingering and painful death – except bizarrely, when one has reason to believe that this is just what will happen. Perhaps this feature of belief-ascription arises because explanation involves locating salient features. It constitutes a part of the truth in externalism concerning belief – where ‘external’ is defined as ‘outside the epistemic perspective of the subject’.

It would certainly be misleading to characterise the product of inference as an unconscious belief. If it is a belief of mine, it is an ordinary one. It is of the nature of beliefs that sometimes they are considered or publicly expressed, and at other times not; this does not imply two categories of belief, conscious and unconscious. Even less plausible is the claim that there could be a conflict between avowed beliefs and non-avowed products of inference. Brandom for instance claims an ambiguity in the concept of belief between what one is prepared to avow, and what one ought rationally to believe, as a consequence of what one is prepared to avow.<sup>17</sup> But convincing examples of conflict are hard to come by. In fact, insofar as advocates of the dispositional model support it with examples of unconscious beliefs – which often they do not – these examples are quite implausible.

One instructive illustration, though it does not involve an inference, is provided by Shoemaker. Before washing the dishes I put my watch in my pocket. Later I look all over the place for it, before I remember where it is; but ‘the information that it was in my pocket was in my memory the whole time, and in some sense I believed it’. It is even possible, Shoemaker maintains, that I would not assent to the possibility that the watch was in my pocket if it was put to me while searching – ‘Yet in a good sense I do believe it’. The same is true, he claims, of many cases of beliefs stored in my memory.<sup>18</sup> Now this is a paradigm example of the reification of belief as ‘stored information’, to the neglect of its normative aspects, notably commitment to the truth of *p*. Where self-deception is not involved, there is no plausibility in the suggestion that I believe what I sincerely disavow.

An important motivation for the idea of unconscious beliefs is the view that every regularity in human behaviour is explicable by a belief. Say it is pointed out to me that I am very careful making right-hand turns, and much less so in my left-hand ones. This does not imply that I *believe* that right-hand turns are, say, more dangerous. Similarly, I may twiddle my moustache when bored, or deep in thought, but if it is a revelation to me that I do so, I can hardly go on to wonder what germane means-end belief rationalises or explains the twiddling. This is simply habitual behaviour. It can be a revelation that I engage in such behaviour; but it can never be a surprise to me that I have a certain belief, any more than that I have a certain intention.

(3) *Previously-formed beliefs*. There are very many cases where I have already thought about the question and already have the appropriate belief. In such cases, if someone asks me whether I believe that *p*, and I can remember having made up my mind, then I avow the belief spontaneously. This could be an

habitual re-statement of the belief. Alternatively, the question may prompt reconsideration, and perhaps a considered restatement. But again there should be no suggestion of a distinct category of unconscious belief, revealed by introspection, or observation of my behaviour.

There is, however, a sense in which previously-formed beliefs can have an external existence. If someone asks me in ten years time – or even in six months – ‘What do you believe about the avowal of belief?’, I will very likely consult ‘The Authority of Avowals and the Concept of Belief’. I might ask myself ‘That’s what I said then, but what do I believe now?’ – a belief that *p* implies a commitment to the truth of *p*, but not an indefinite one. But I may simply reaffirm my commitment to the beliefs expressed therein. This would be a case where I can articulate my beliefs only with the help of an external aid, and so in a limited way I adopt a third-person perspective on my own beliefs – in the sense in which I can discover someone else’s beliefs by looking at what they have written. But there are crucial differences. I have to have confidence that I did a thorough job in the article; and it has to be a reasonably live issue for me. For example, if someone asks me what I believe about the causes of the Russian Revolution, it would be regarded as a joke – or an expression of bizarre egotism – if I brought out a school essay of mine which I have never looked at since, and said ‘Read this’.

In many cases avowals of belief are merely habitual, or unreflective; or reflection is inappropriate or impossible. Does this suggest restrictions on the decision principle? The Creature of Habit repeatedly says ‘I always say that the A68 is a much safer road than the A1’. He or she will not ask the question ‘That’s what I said then, but what do I believe now?’ This person’s beliefs, once formed, are highly recalcitrant to revision. The Dogmatist is a slightly more reflective version of the Creature of Habit. Both are happy with the beliefs inherited from their parents or upbringing, or those acquired long ago. But Dogmatists, unlike Creatures of Habit, are able to articulate their grounds for believing that *p*, for instance ‘Because I was taught that it was so’ or ‘Because I’ve always believed it’. We are all Dogmatists on certain questions. Concerning so-called groundless beliefs, there is no decision procedure for *p*, at least for the subject in question at the time in question. Among these are religious beliefs, ethical principles, and other basic beliefs such as ‘The earth is round’ and ‘I have two arms’. (Much more could be said about this important category.)

It may be suggested that for Dogmatists, the basis for present-tense avowal – apparently, one’s past decisions or behaviour – is symmetric with that for third-person ascription. The implication is that the decision principle applies only to the autonomous, rational enquirer. But reference to one’s past decisions or behaviour, even for the Dogmatist, occupies a different role in the first- and third-person cases. If I am a Dogmatist, any work which I at one time put into deciding whether *p*, now stands as grounds for my continued commitment to the truth of *p*. It is only derivatively, via the decision principle, that it stands as grounds for the avowal of the *belief* that *p*; that avowal has no independent support. No alternative basis for self-ascription, and no distinctive possibility of error, results.



Even for Dogmatists, avowals constitute a commitment by the ascriber to the truth of  $p$ , in a way that third-person ascriptions do not.

(4) *Mere responses to perceptual information.* When driving, I move out to avoid a parked car; so, it may be said, I believed there was a parked car in my path. But if the perception is subliminal, I may not avow the belief. I would argue, however, that such cases involve, at best, an informational state more primitive than belief. Explanation in terms of a practical inference schema – in terms of reasons for action – is not appropriate at this level. The status of informational states is addressed further below.

I said earlier that the alleged holism of the mental seems to conflict with the authority of avowals. It also seems in conflict with the decision principle. Davidson's theory of interpretation has a tendency to portray my question, in deciding whether  $p$ , as 'What must I say about  $p$ , in order to maintain a coherent set of beliefs and thus a plausible interpretation of my own behaviour?' This portrayal is mistaken.

Say that at one time I avowed the belief that  $p$ , and now avow the belief that  $q$ . Someone points out that  $p$  and  $q$  are inconsistent. If I come to agree with them, I will alter my belief on either  $p$  or  $q$ . I may be a Dogmatist, loathe to admit I am ever wrong; this will influence my response to the critic's assertion 'But  $p$  implies not- $q$ '. But that response is to the questions whether  $p$ , whether  $q$ , and whether  $p$  implies  $q$ ; not to whether I believe that  $p$ , etc. The tension is not: I avowed  $p$  at  $t$ , and  $q$  at  $t+n$ , and these are inconsistent, and so my *behaviour* is not rationally interpretable; rather, it is my *beliefs* that are not coherent. The desire, or requirement, for coherence in one's beliefs does not suggest a behavioural basis for their self-ascription; nor does it imply a compromise of the authority principle.<sup>19</sup>

### 3. Expression and Non-Cognitivism: Resolving the Dispositional Dilemma

If the preceding defence of the authority and decision principles is felt to carry conviction, reliabilism must be rejected as a resolution of the dispositional dilemma. I will now defend an alternative, non-cognitive position that undercuts the traditional debate. On this view, the answer to the question of how I could come to have apparently immediate and certain knowledge of a state extended in time is simply: 'I do not'. Avowals of belief, I will argue, are *expressions* and not descriptions of one's beliefs; the self-scanning, descriptive or reporting model is rejected. Avowals should be assimilated to assertions of the fact believed, in the sense suggested by what may be termed the *assertion thesis*: 'I believe that  $p$ ' functions as a (sometimes tentative) assertion of  $p$ .<sup>20</sup> It is not possible for the subject merely to report their belief that  $p$ , without taking a stand on, or expressing a commitment to, the truth of  $p$ . Avowals of degrees of belief should similarly be regarded as expressions and not descriptions of the degree

of belief. 'I think I believe . . .' expresses hesitancy over whether  $p$ , not over whether I believe that  $p$ .

The claims in the preceding paragraph add up to a package of views, some more problematic than others, and none of them widely held. I will move from discussion of expression to the more crucial non-cognitive thesis. Although for obvious reasons the present treatment has focussed on avowals of belief, the primary expression of belief is the straight assertion of the embedded proposition – otherwise there would be little point in an assertion thesis which assimilates avowals to it. Assertion and belief comprise a circle of concepts. It is not possible to define assertion in terms of some allegedly more basic, independent notion of belief; nor to define belief in terms of some more basic, independent notion of disposition to assert. The assertion thesis assumes the conceptual inter-dependence of belief and assertion in at least this sense: one cannot report a present belief without also asserting it.

The claim that avowals are expressive contrasts with a less plausible way of rejecting the self-scanning or reporting model, viz. a performative analysis. Both expressive and performative accounts assert a conceptual connection between assertion and belief, denying the functionalist view that the avowal is a causal consequence of the belief. But a performative analysis seems indefensible. Whenever I sincerely, etc., avow 'I believe that  $p$ ', it is the case that I believe that  $p$ ; but the utterance of the avowal does not *make it* the case that I believe that  $p$ . In contrast, although there are conditions which have to be satisfied, it is much more plausible to say that the utterance 'I promise . . .' makes it the case that I have promised. The latter claim is reinforced when one considers that sincerity is not required. A lying promise is still a promise; though if the words are uttered by a compulsive confabulator, one might hold that no promise had been made because they did not really understand what they were saying. These would be two different senses of the speaker 'not meaning' what they had said.<sup>21</sup>

So 'I believe that  $p$ ' expresses the belief that  $p$ ; 'I promise . . .' does not express a promise, it makes one. But what does the claim of expression amount to? Certainly, in the case of expressive behaviour, one does not infer from the behaviour to the existence of the pain or belief. Where behaviour counts as an expression of what is ascribed, the question 'Why does what is ascribed have that kind of manifestation?' is foolish or has no answer. But expressive behaviour is not just behaviour that is a criterion – 'necessarily good evidence' – for belief.

What more is it then? The concept remains frustratingly elusive, hence perhaps the paucity of discussion of it in the philosophy of mind. Attempts to apply it in the context of avowals confront the objection that 'express' is being applied in two quite separate senses to non-verbal and verbal behaviour, resulting in an equivocation. The interesting sense, it may be said, is not applicable to the case of belief – there is no primitive expression, no natural, pre-linguistic belief-behaviour analogous to that of pain, of primitive emotions such as fear, and perhaps desire. But in the banal sense of 'give utterance to' or 'put into words', it is simply a truism that the avowal is an expression of belief, and therefore readily accepted by a

descriptive account. So, it is argued, it is either false or innocuous to say that avowals of belief are expressive.

Now there are cases where the concept of expression equally implies verbal and non-verbal behaviour, even if the latter is not primitively expressive. There is a unified sense of expression in the case of non-primitive feelings – notably those which are conceptually too complex to be ascribed to animals, such as remorse and indignation – which involves both putting into words and characteristic non-verbal behaviour. To say that someone is not good at expressing their feelings, is to say they are inadequate both in putting them into words, and in their non-verbal demonstration of affection, anger, etc. It hardly means: not very good at naturally expressing primitive feelings.

However, this unified sense of ‘express’ does not seem to apply in the case of belief. Feelings are the kinds of thing which naturally tend to be expressed – we talk of repressing or bottling them up. The same is perhaps true of opinions, views and convictions, but more mundane beliefs seem ‘inert’. In the case of belief, ‘express’ does seem to mean simply ‘give utterance to’. It is not plausible to regard non-verbal behaviour as expressive of belief. ‘She expressed her belief that East Timor is a dangerous place by not going there’ sounds absurd.

It is useful to explore a more prosaic example. It is unlikely that someone’s waiting at the bus-stop could be said to express their belief that the bus is coming soon. If it did, it equally expresses their desire to catch the bus, and their intention to do so. The expressive claim is even less plausible when the situation is a little less straightforward. Say the subject believes that someone they wish to avoid will be on the bus, so they keep away from the bus-stop when the bus is due to arrive. Here the ascription of the belief that the bus is coming soon seems to be inferential, resting on the supposition that the subject believes that someone they wish to avoid is on the bus. It may also be inferential in the simpler case. (Uncertainty results from the problematic characterisation of ‘inferential’.)

But it is not because of the presence of belief-desire holism that non-linguistic behaviour fails to be expressive of belief. Rather it is because the behaviour reveals the belief in no stronger sense than simply letting it be seen. In contrast, an avowal or straight assertion reveals the belief by communicating it. And ‘giving utterance to’ – communicating or saying – is not a banal sense of ‘express’. It is at the core of expression in all its senses.<sup>22</sup> The assertion and the avowal therefore occupy a privileged position in the concept of belief.

Like the concept of self-consciousness itself, expression is a theoretical notion with roots in ordinary use. Its theoretical nature arises from the conceptual role of belief, intention and desire in the constitution of a subject of thoughts and experiences. It would be an attenuated subject which could simply report its beliefs, objectively and from a detached viewpoint, as if they hardly belonged to it – that is, with no commitment to their truth. Neglect of such considerations, and an exclusive focus on the pain paradigm of expression, lies behind the standard assumption that an expressive thesis implies a non-assertoric thesis – one which denies that avowals are truth-assessable. Now it could be argued that – its label notwithstanding – the assertion thesis gives a non-assertoric account of, say, ‘I

believe that it is raining', one according to which the utterance does not assert that I believe that it is raining. But what is the difference between asserting that I believe that it's raining, and asserting that it's raining, other than a possible tentativeness in the former case?

The assumption that expressive implies non-assertoric is found in Crispin Wright's recent discussion, where he argues that 'the correct explanation of the [authority, transparency and groundlessness of] avowals cannot have anything to do with illocutionary distinctions'.<sup>23</sup> This is surely correct. The proper conclusion to draw, however, is not that expression has no interesting role to play in the treatment of avowals – but rather, that the distinction between expression and description is not one within speech-act theory. It is a more fundamental epistemic distinction which draws on the considerations advanced in the preceding discussion.

It would be good if the concept of expression could be made more precise.<sup>24</sup> But it is the non-cognitive thesis which is at the heart of the resolution of the dispositional dilemma. Its claim is that 'I believe that/suspect that/know that/am of the opinion that . . . I believe that  $p$ ' has no serious cognitive use. This is one important reason why Wright's view that intentional states are constituted by the subject's 'best opinion' is misguided; I cannot be 'of the opinion' that I believe that  $p$ , nor indeed that I intend to do  $A$ .<sup>25</sup>

'I know that I believe that  $p$ ' does not express a genuine knowledge-claim. It may have the force of 'I concede . . .' in 'I know I believe that capital punishment is wrong, but that does not mean I am against deterrent sentences'; but this is just equivalent to 'Though I believe that . . . this doesn't mean . ..'. 'I don't know what I believe' can only mean 'I've not made up my mind on this question'; it cannot mean 'I have beliefs on this question but I don't know what they are'. It is not only the explicitly first-person cases which are ruled out by the non-cognitive thesis. What is rejected is any instance of 'X believes that X believes that  $p$ ' where X refers to the same subject, a subject who allegedly at some time has a present belief concerning a present belief. In contrast, there is clearly no problem with 'AH believes that EJJ believes that . . .', and so on ad infinitum; or with 'I believe that I believed that  $p$ '.

Remarkably, even advocates of the authority of avowals see no problem with the idea that one can have beliefs about ones own beliefs.<sup>26</sup> It is assumed that because 'I believe that  $p$ ' is a proposition, it can meaningfully be substituted for  $p$  in the schema 'I believe that  $p$ '. This no more follows, than it follows from the logically non-contradictory status of the conjunction of ' $p$ ' and 'I believe that not- $p$ ', that one can meaningfully assert ' $p$ , but I believe that not- $p$ ' (the issue of Moore's Paradox).<sup>27</sup> As Wittgenstein's remarks suggest, the claim of uniform substitution rests on a narrow formalism (substitute 'I believe that I believe that  $p$ ' for  $\sqrt{-1}$ , and 'I believed that I believe that  $p$ ' for  $\sqrt{1}$ ):

'But surely "I believed" must tell of the same thing in the past as "I believe" in the present!' – Surely  $\sqrt{-1}$  must mean just the same in relation to  $-1$ , as  $\sqrt{1}$  means in relation to  $1$ ! This means nothing at all'.<sup>28</sup>

The lack of intelligible actual examples suggests that opponents of a non-cognitive thesis cannot explain what 'I believe that I believe that  $p$ ' is supposed to mean. It takes its sense entirely from the third-person or mixed case ('I believe that Jonathan believes that I believe . . .' etc.), or from cases involving other tenses ('I believe that I believed that  $p$ '). The idea that the knowledge-claim makes sense derives from the theoretical assumption that belief is a disposition, whose existence is reported by the subject in their self-ascriptions.

Non-cognitivism has a close affinity with the assertion thesis, which argues against the too-ready assumption that something other than 'It's raining' may be asserted by 'I believe that it's raining'. The decision principle also makes a non-cognitive account hard to avoid. If, having decided the question whether  $p$ , there is no further question whether I believe that  $p$ , it is difficult to see how the avowal could express a possible item of knowledge for me – viz. knowledge of a present belief. Any attempt to assess the truth of 'I believe that  $p$ ', as distinct from the truth of the embedded proposition, is a charade.

Rejection of cognitivism goes together with rejection of the reporting model. The idea that I observe, from an external standpoint as it were, my present beliefs, without expressing a commitment to them but merely expressing a belief that I have them, suggests a dissociated personality. It has a flavour of the psychotic. The idea of non-avowed present beliefs is analogous to the thought-insertions or auditory hallucinations experienced by schizophrenics – thoughts had by the subject yet felt not to belong to them. But beliefs are not things that just come into my head, that I can observe with concern, fear or approval. I can take a view about  $p$ , but not about my present belief concerning  $p$ ; my present beliefs precisely embody my views.

#### 4. The Concept of Belief

The dispositional model not only gets the first-person epistemology wrong; it also fails to explain the function of the avowal and the straight assertion in third-person ascription, and hence the explanatory role of belief. Even in the third-person case, belief does not have the character of a disposition. It would thus be quite wrong to take the authority of avowals as implying that belief that  $p$  is constituted by a disposition to assert that  $p$ . Our assertions are, in at least a rough and ready way, composed to fit the occasion, in contrast with true dispositions to verbal behaviour such as: 'When I feel hostility from my audience, I am disposed to use long words and say things like *mutatis mutandis* a lot'. Or: 'When I feel I am being dismissed as a dull academic, I find myself making exaggerated claims about my experiences in a rock 'n' roll band'.<sup>29</sup> This behaviour is largely involuntary, and the dispositions are learned about inferentially, in a third-personal way.

The dispositional model rests on a prevalent but simplistic divide between conscious episodes and dispositions. 'Disposition or episode?' is as crude and

arbitrary as 'Physical or mental?'. Paradigm psychological or behavioural dispositions are dispositions to drive over the speed limit, to smoke cigarettes when anxious or to use long words when feeling threatened by an audience. Character- and personality-traits such as generosity and obstinacy are other, perhaps less clear, examples.<sup>30</sup> In the paradigm cases there is no authority, indeed no avowal in the ordinary sense. Someone can smoke heavily when anxious or be obstinate without it ever occurring to them that they are, and without admitting it when it is put to them. This need involve no cognitive defect in the subject in the way that self-deception does. In contrast to avowals of belief, self-ascriptions of obstinacy merely report, reliably at best, the existence of the state.

Belief is neither a disposition nor an episode. It should be regarded instead as one of a range of psychological concepts which exhibit different biases towards one or other end of the episode-disposition continuum. Expecting, hoping, intending and wishing – though probably not desiring or wanting – are other examples. Each of these possesses its own version of the authority and decision principles. Beliefs, intentions, expectations and so on explain and justify dispositions to behave and so cannot simply consist in them. Ascribing a belief rationally explains the behaviour, it makes it apt or sensible rather than merely habitual. Functionalism, in postulating an inner state that causes behaviour, attempts to acknowledge the explanatory role of belief, but in a mechanistic form which cannot accommodate the authority of avowals. The framework within which the concept of belief operates is not simply a causal nexus but a rational, interpretative one involving the authority of avowals. Beliefs belong to the sphere of reasons.

Dispositional or informational state accounts in general fail to capture the justificatory role or normativity of belief-ascription. Ordinary psychological dispositions lack anything analogous to the *commitment to the truth* or *aiming at the truth* which is essential to belief. The nature of such commitment is highly elusive, and a proper account of it would bring in agency, rationality, knowledge, self-consciousness and much else besides. But to acknowledge it, reference to judgment and avowal is essential. Certainly commitment to truth cannot be defined independently of belief or its cognates; 'commitment' and 'acceptance' are not more basic than 'belief'. It follows that Peacocke is mistaken in suggesting that 'representing a content', and 'accepting' it, can feature in a non-circular possession condition for belief. Collins' non-constitutive account of belief, with which I have much sympathy, also seems to regard 'commitment' as more basic than 'belief', neglecting its connection to behaviour – to dispositions and avowals.<sup>31</sup>

Commitment to truth is a hallmark of beliefs but not of informational states. These are meant to be passive, not conscious. As Gareth Evans put it, they are states of the informational system which is the substratum of our cognitive lives. His position is that belief is defined – similarly to the present account – in terms of judgment and reasons, while 'informational state' should be defined independently as a primitive and fundamental notion. But once it is recognised that the basic notion of information is personal, not impersonal – one which implies a

recognised way for the subject to gain knowledge by means of it – then it becomes clear that if it has application at all, the concept of ‘informational state’ must be derived from that of belief and not vice versa.<sup>32</sup>

In conclusion it is necessary to return to the status of the authority principle, particularly in the light of externalism. Proponents of Davidsonian interpretation emphasise that the ascription of a sincere, comprehending avowal is itself the product of interpretation of the subject’s beliefs and desires; so the principle seems not to provide an effective decision procedure for the ascription of belief. Now although the authority principle is a regulative one, ruling out the possibility of honest error, there is still a fact of the matter about whether a subject makes a sincere avowal – at least in that it makes no sense to suppose that someone can lie without knowing it. Deceit requires the intention to deceive, and a subject’s avowals of intention are authoritative. So ‘I am sincere in my avowal’ is itself authoritative.

The subject’s lack of authority concerning what they mean may still occasion doubts about the authority principle. Now in fact there is limited authority here. What I mean on a particular occasion is what I intend to say, and so is subsumed under the authority of avowals of intention. Thus there is an authority concerning the speaker’s meaning. Whether the intention is realised is another matter; linguistic meaning is a social construct concerning which the subject cannot make authoritative avowals. The linguistic meaning of ‘It’s getting cold in here’ is given by disquotation and authority is not exhibited; concerning the speaker’s meaning, which might be expressed as ‘Please shut the door!’, the subject’s avowal is authoritative.

The incompleteness of this authority is latched onto by many proponents of externalism. It is not clear how far an externalist critique of authority has to assume a reporting model of avowals; certainly once it is allowed that one may have thoughts or opinions about one’s present belief, it seems possible that such thoughts or opinions can be mistaken. However, rejection of the reporting model means that one popular deflationary response to externalist objections is not available; viz. that *p* has the same content in ‘I believe that I believe that *p*’ as in ‘I believe that *p*’. The authority of avowals must not be formulated in terms of present beliefs about present beliefs; as defined earlier it makes reference to the corresponding third-person ascription.

A defender of authority has to suppose that *p* has the same sense in ‘I believe that *p*’, and ‘X believes that *p*’. A comprehending avowal and constancy of sense across the first- and third-person ascriptions was in fact specified in the authority principle as defined earlier. But this may seem to grant too easy a victory over externalist critics. What guarantee of sameness of content obtains in an ordinary communicative situation, when the speaker’s words are transposed into the third-person? One can say only that both subjects apparently speak the same language and understand what each other is saying. One could also ask what guarantee there is that *p* has the same sense in ‘I believe that *p*’ and in ‘I believe that I believe that *p*’ uttered by the same person a few seconds later. (If it is helpful to do so, the authority principle might be re-formulated without

reference to the third-person ascription, in terms of the subject's immunity to error.)

The debate is increasingly arcane and too many trees have been felled in its cause. But I am inclined to say that authority and externalism are incompatible. If externalism does have the absurd consequence that I do not really know what I think or believe, so much the worse for externalism. However, it is instructive to conclude by commenting briefly on a pervasive if weak variety of non-comprehension in avowals which is raised by Kripke's celebrated puzzle about belief.<sup>33</sup> The case of Pierre should be familiar given the literature it has generated. The external fact of which Pierre is unaware is that London = Londres. If Pierre avows 'I believe that London is not pretty', one might hesitate to ascribe that belief, given that he is also prepared to avow 'Londres est jolie'. As a result the authority of avowals of belief about London and Londres seems called into question.

As Kripke recognised, this phenomenon is found within as well as across linguistic communities. Ignorance of the identity Mark Sinker = Hopey Glass could lead to the expression of contradictory beliefs, for instance 'Mark Sinker is editor of "The Wire"' and 'It's not the case that Hopey Glass is editor of "The Wire"'. Presumably contradictory beliefs of this kind are quite prevalent. The subject's belief would become irrational only if, on being told and apparently acknowledging that MS = HG, they persist in their original claims.

There are many possible avenues of resolution to the puzzle. To return to Kripke's original example, the problem seems to be: Just what has Pierre avowed – which proposition or content? One does not want to say that the objects of belief are sentences-in-a-language; but the idea that its objects are propositions is under strain. A deflationary resolution, comparable to that for self-deception, is preferable, although in the present case the subject is not guilty of irrationality or illogicality. It allows the possibility, too readily rejected by most commentators, that contradictory beliefs may be attributed to a rational agent.<sup>34</sup> Kripke's story is a perfectly intelligible one, and having told it there is nothing more to say concerning what Pierre believes. The further question 'So which proposition exactly does he believe?' should be resisted.

On this deflationary position, the consequences of the puzzle for the authority principle will not be dramatic. Say I hear Pierre's avowal 'London is pretty' and ascribe that belief to him. The ascription is not a false one, but if I subsequently learn the full story I will realise that it is an incomplete answer to the question 'What does Pierre believe concerning the question whether London is pretty?' That is all.

The debate over externalism and authority has tended to obscure the central features of the concept of belief. In this article I have tried to cast light on some of those features by reviving the insights of mental act accounts of belief, whilst avoiding their Cartesian implications. I have suggested that treatments of belief which neglect the authority of avowals, and aspects associated with it, can yield only an impoverished account of the subject of thoughts and experiences. At the heart of this treatment is the elusive notion of expression, on



which much more work remains to be done. But that is material for another occasion.<sup>35</sup>

Andy Hamilton  
 Department of Philosophy  
 Durham University  
 Durham DH1 3HP  
 England  
 A.J.Hamilton@durham.ac.uk

## NOTES

<sup>1</sup> For instance Russell 1921, Ch. XII; Ramsay 1990, especially p. 40.

<sup>2</sup> Dennett 1981: 306–7. Dennett's distinction between 'opinion' and 'belief' contrasts with the distinction among philosophers of science between 'acceptance' – which seems to mean belief caused by a commitment – and belief based on evidence.

<sup>3</sup> Endorsed for instance by Pears 1984: 50.

<sup>4</sup> For instance 'I believe that you are a philosopher' would be transformed into 'X believes that [the person addressed] is a philosopher'.

<sup>5</sup> Evans 1982: 225. As Evans notes, the principle originates with Wittgenstein. Gallois properly claims that for me, the question whether *p* is not distinguishable from the question whether I believe that *p*. But he is wrong to treat this as involving a justified *inference* from '*p*' to 'I believe that *p*' – this is surely otiose (Gallois 1996).

<sup>6</sup> Something like the Authority Principle has been defended by Phillips Griffiths (1967), Burge (1988), Wright *et al.* eds (1998, 'Introduction'), and Bilgrami (1998).

<sup>7</sup> A subtle treatment of this issue, from the viewpoint of a kind of externalism, is found in Bilgrami 1992: 49–56. It is also discussed in Wright *et al.* (eds.), 1998: 7–10 and *passim*.

<sup>8</sup> The more recent definition 'inside the person's skin', used by Boghossian (1998) and Davies (1998), is no better.

<sup>9</sup> This dilemma has been noted by various writers, including Wright (1989: 237), Bilgrami (1992: 247–8), and Jacobsen (1996).

<sup>10</sup> Davidson 1987. See also his 1984 and 1989.

<sup>11</sup> Peacocke 1992: 155.

<sup>12</sup> See for instance Davidson 1985.

<sup>13</sup> See for instance Bittner 1988 and Darwall 1988.

<sup>14</sup> Some would argue that self-deception is not wholly irrational, since the subject is attempting to maintain their emotional equilibrium in the face of a painful situation. Gallois for instance claims self-deceivers are deploying strategic as opposed to cognitive reasons (1996: 103–6).

<sup>15</sup> Bilgrami (1998) maintains that self-deception is not a threat to the authority of avowals, while still assuming that the subject believes both that *p* and that not-*p*. This seems an inaccurate picture of self-deception and also too sanguine about its effect on authority.

<sup>16</sup> This is the view of Dummett 1981: 286–7.

<sup>17</sup> Brandom 1994: 195–6.

<sup>18</sup> Shoemaker 1996: 80.

<sup>19</sup> Other criticisms of Davidson's account are found in Moran 1994, and Hacker 1997.

<sup>20</sup> The *locus classicus* of the assertion thesis is Wittgenstein 1958: 190–2. The assertion thesis constitutes one part of an explanation of Moore's Paradox. It has been powerfully defended by Collins 1987, to which I am much indebted.

<sup>21</sup> Jane Heal has attempted a performative analysis of avowals of belief in her unpublished 'First-Person Authority', by dint of also discerning a self-descriptive element in orthodox performatives such as promising.

<sup>22</sup> The point is made in Charles Taylor 1979, the best discussion of expression in the literature. He regards communication, in the sense in which smiling communicates friendliness and does not simply allow it to be seen, as central to expression. Although indebted to his account, I disagree with his claim that what is expressed can be manifest only in some expression or other.

<sup>23</sup> Wright 1998: 38; the assumption is also made by Jacobsen in his subtly-argued 1996. These interpretations are perhaps also influenced by the status of ethical emotivism as a classic 'expressive theory'.

<sup>24</sup> The issue is discussed further in Hamilton (forthcoming).

<sup>25</sup> Wright 1989: 250–54.

<sup>26</sup> For instance, none of the authors in Wright *et al.* (eds.) (1998) questions it.

<sup>27</sup> The analogy is not perfect. Moore's Paradox arises because one can for instance entertain the hypothesis that I believe that *p* but not-*p*; one cannot so much as entertain the hypothesis that I believe that I believe that *p*. The substitutivity objection was put to me by John McDowell. (Collins 1996 gives an analysis of Moore's Paradox which in key respects agrees with the present account of belief.)

<sup>28</sup> Wittgenstein 1958: Part II, p. 190.

<sup>29</sup> I am grateful to Arthur Collins for this point.

<sup>30</sup> In his discussion of 'Jones was brave', Dummett rightly refers to our residual commitment to character as a 'spiritual mechanism', in place before its effects are displayed in behaviour (Dummett 1978: 15); but perhaps this is simply a categorical basis to the disposition.

<sup>31</sup> Peacocke 1992, Ch. 6; Collins 1987, Ch. II. Brandom argues that, rather than offering an analysis of belief in terms of commitment, one should replace it with a technical concept of doxastic commitment (Brandom 1994: 195–6). An heroic attempt to link the normativity of intentional states to agency and responsibility is found in Bilgrami 1998.

<sup>32</sup> Evans 1982: 123–4. The impersonal concept of information is criticised in Hamilton 1995, and in Hamilton (forthcoming). It is therefore an attenuated concept of belief that is applicable to the higher mammals, not a fundamental concept of informational state applicable also to humans – and certainly not that continuity between human and animal beliefs secured, at unacceptable cost, by the dispositional model. Clearly much more would need to be said, on another occasion, concerning animal beliefs.

<sup>33</sup> Kripke 1975: 239–283.

<sup>34</sup> This is a possibility neglected by the subtle and complex treatment found in Moore 1999.

<sup>35</sup> I am very grateful for discussions with Arthur Collins, Robin Hendry, E.J. Lowe, Roger Squires, Crispin Wright, and audiences at Research Seminars at Keele, Lancaster, Leeds and Munich. Also for comments from the anonymous referees of this journal.

## REFERENCES

- Armstrong, D. (1973), *Belief, Truth and Knowledge*. Cambridge: Cambridge University Press.  
Bilgrami, A. (1992), *Belief and Meaning*. Oxford: Blackwell.

- Bilgrami, A (1998), 'Self-Knowledge and Resentment', in Wright *et al.* (eds.) 1998.
- Bittner, R. (1988), 'Understanding a Self-Deceiver', in McLaughlin & Rorty (eds.) 1988.
- Blackburn, S. (1984), *Spreading The Word*. Oxford: Clarendon.
- Bradt, J., ed. (1989), *The Mind of Donald Davidson*. Amsterdam: Rodopi.
- Brand, R. (1994), *Making It Explicit: Reasoning, Representing and Discursive Commitment*. Cambridge, MA: Harvard University Press.
- Burge, T. (1988), 'Individualism and Self-Knowledge', in *Journal of Philosophy* 85, 649–665.
- Cassam, Q., ed. (1994), *Self-Knowledge*. Oxford: University Press.
- Collins, A. (1987), *The Nature of Mental Things*. Indiana: University of Notre Dame Press.
- Collins, A. (1996), 'Moore's Paradox and Epistemic Risk', in *Philosophical Quarterly*, 46, 308–19.
- Darwall, S. (1988), 'Self-Deception, Autonomy, and Moral Constitution', in McLaughlin and Rorty (eds.) 1988.
- Davidson, D. (1984), 'First Person Authority', in *Dialectica*, 38, 101–11.
- Davidson, D. (1985), 'Deception and Division', in Le Pore and McLaughlin (eds.) 1985.
- Davidson, D. (1987), 'Knowing One's Own Mind', in *Proceedings and Addresses of the American Philosophical Association*, 60, 441–58; also in Cassam (ed.) 1994.
- Davidson, D. (1989), 'What Is Present To The Mind', in Bradt (ed.) 1989.
- Dennett, D. (1981), *Brainstorms*. Brighton: Harvester.
- Dummett, M. (1978), 'Truth', in his *Truth and Other Enigmas*, London: Duckworth.
- Dummett, M. (1981), *Frege: Philosophy of Language* (2nd edition). London: Duckworth.
- Evans, G. (1982), *The Varieties of Reference*. Oxford: Clarendon.
- Gallois, A. (1996), *The World Without, the Mind Within: An Essay on First-Person Authority*. Cambridge: University Press.
- George, A., ed. (1989), *Reflections on Chomsky*. Oxford: Blackwell.
- Hacker, P.M.S. (1997), 'Davidson on First-Person Authority', in *Philosophical Quarterly*, 47, 285–304.
- Hamilton, A. (1995), 'A New Look at Personal Identity', in *Philosophical Quarterly*, 45, 332–49.
- Hamilton, A. (forthcoming), *Memory and the Body: A Study of Self-consciousness*.
- Heal, J. (unpublished), 'First-Person Authority'.
- Jacobsen, R. (1996), 'Wittgenstein on Self-Knowledge and Self-Expression', in *Philosophical Quarterly*, 46, 12–30.
- Kripke, S. (1975), 'A Puzzle About Belief', in Margalit (ed.) 1975.
- Le Pore, E. and McLaughlin, B., eds. (1985), *Actions and Events*. Oxford: Blackwell.
- Margalit, A., ed. (1975), *Meaning and Use*. Dordrecht: Reidel.
- McLaughlin, B. and Rorty, A., eds. (1988), *Perspectives on Self-Deception*. Berkeley: University of California Press.
- Moore, J. (1999), 'Misdisquotation and Substitutivity: When Not to Infer Belief From Assent', in *Mind* 108, 335–65.
- Moran, R. (1994), 'Interpretation Theory and the First-Person', in *Philosophical Quarterly* 44, 154–173.
- Peacocke, C. (1992), *A Study of Concepts*. Cambridge, MA: MIT Press.
- Pears, D. (1984), *Motivated Irrationality*. Oxford: Clarendon.
- Phillips Griffiths, A. (1967), 'On Belief', in Phillips Griffiths (ed.) 1967.
- Phillips Griffiths, A., ed. (1967), *Knowledge and Belief*. Oxford: University Press.
- Ramsay, F. (1990), 'Facts and Propositions', in H. Mellor (ed.) *Philosophical Papers*. Cambridge: Cambridge University Press.
- Russell, B. (1921), *The Analysis of Mind*. London: George Allen and Unwin.

- Shoemaker, S. (1996), 'Moore's Paradox and Self-Knowledge', in his *The First-Person Perspective And Other Essays*. Cambridge: Cambridge University Press.
- Taylor, C. (1979), 'Action As Expression' in C. Diamond and J. Teichman (eds.) *Intention and Intentionality*. Brighton: Harvester.
- Wittgenstein, L. (1958), *Philosophical Investigations*. Oxford: Blackwell.
- Wright, C. (1989), 'Wittgenstein's Rule-Following Considerations and the Central Project of Theoretical Linguistics', in George (ed.) 1989.
- Wright, C. (1992), *Truth and Objectivity*. Cambridge, MA: Harvard University Press.
- Wright, C. (1998) 'Self-Knowledge: The Wittgensteinian Legacy', in Wright *et al.* (eds.) 1998.
- Wright, C., Macdonald C., and Smith, B., eds. (1998), *Knowing Our Own Minds*. Oxford: Oxford University Press.