Abstracta 3 : 1 pp. 46 – 73, 2006

GULLIBLE YET INTELLIGIBLE

Daniel O'Brien

Abstract
In this paper I describe the imaginary community of Gullible. Gulliblians are led by moral pressures to believe whatever they are told and, in the scenario that I sketch, this leads to them having widespread contradictory beliefs. This community is nevertheless intelligible to us given what we know about their situation and their moral code. Davidson, however, holds there to be what I call a logicist constraint on interpretation: thinkers can only be interpreted if a good proportion of their beliefs are rational and, for Davidson, rationality entails logical consistency. The possibility of Gullible therefore forces us to reject this logicist aspect of Davidson's account of interpretation and intelligibility.

This paper argues for a pragmatist account of intelligibility, one that is at odds with Davidson's account of interpretation. Davidson argues that logically inconsistent beliefs can only be made sense of against a background of mainly consistent beliefs. I argue that this is not so. Logical consistency is not necessary for intelligible thought. In section 1, I explain Davidson's interpretationist approach and, in section 2, I note that he is also committed to what I call logicism, the view that in order to interpret thinkers we must assume that they do not hold inconsistent beliefs. Section 3 focuses on the empathetic aspect of interpretation. In section 4, I present my key thought experiment, that of a community called Gullible in which inconsistent beliefs are rife. Section 5 provides moral, psychological and pragmatic considerations to support the claim that Gullible is an intelligible community and, in section 6, I reject various lines of argument that question the coherence of the Gullible scenario. Section 7 notes that Davidson does allow some inconsistent beliefs, but he is nevertheless driven by unwarranted logicist assumptions; and in the conclusion, section 9, I return to empathy and to how this aspect of interpretation constrains what we can come to see as intelligible. To begin, then, let us turn to Davidson and his interpretationist approach.

1. Interpretationism

Interpretationists claim that a study of the practice and assumptions behind interpretation can reveal features integral to the structure of the mind. Thus: 'Interpretationists think that we can gain an understanding of the nature of the mental by reflecting on the nature of interpretation.' (Child, 1994: 6) Our understanding of what it is to have a mind relies on certain facts about us and our interpretative practices. Our behaviour towards the world and each other is regular in various ways, and we make sense of this regularity through interpretation. We account for some of our behaviour in purely physical terms: we fall over cliffs as apples fall off trees, under the influence of gravity. But there are other things that we do that call for the ascription of intentional states, and to make sense of some of our actions we are seen as acting correctly and as acting for reasons. Interpretation picks up on this rational and normative structure. And the interpretationist claims that our conception of a thinker is that of a creature revealed to us through interpretation as rational. The only understanding we have of another having thoughts about the world is one that is constrained by our notion of a thinker being the kind of creature that becomes manifest to us as we practice interpretation.

> [T]here simply is no conception of rational agenthood which is explanatory prior to the interpretationist's account of an agent as a creature which interprets and is interpreted by others; the idea of a rational agent and the idea of rational interpretability arise essentially in the context of communication and interpretation. (Child, 1994: 55)

Such an approach is appealing because an explanation is given of what it is to have a mind that only makes reference to the practice of interpretation. Such a practice involves the observation of a creature's behaviour and thus the philosophy of mind is grounded in something with which we are all familiar; we need not concern ourselves with neurophysiological entities or Cartesian mind stuff.

Davidson is an interpretationist and, for him, to interpret another we have to make certain assumptions about the rationality of our subject: we must assume that she acts for reasons, that her reasons are rationally related to what she believes about the world, and that what she believes is mostly true. These are what Davidson calls the 'principles of charity'. If we do not make such assumptions then interpretation cannot get off the ground. It is important here that the status of these constraints is understood

correctly. It is not that we have an independent conception of what it is to be a thinker and that the application of the principles of charity allows us to narrow down a determinate interpretation of the content of a thinker's thoughts. It is, rather, that the principles of charity are a pre-condition for interpretation to proceed, and thus our conception of what it is to have determinate thoughts about the world is inseparable from the assumed application of just such constraints on interpretation. We shall now go on to look at a key aspect of Davidson's account of the principles of charity, and at why this should be rejected.

2. Logicism

Certain philosophers claim that an account can be given of the essence of rationality in terms of a set of logical constraints that underlie the thoughts of all rational thinkers. I shall call this the 'logicist' conception of rationality. First, logicists take the thoughts of a rational thinker to be constrained by principles of consistency. If it is brought to a thinker's attention that she has acquired the belief that *p* and the belief that not-*p*, then she should decide which one is in fact justified, or decide that she must hold back from either belief until she has further evidence one way or the other. She should not both believe that *p*, and that not-*p*. Further assumptions are also made about how her beliefs are related together through valid inference. If she believes that *p,* and that *p* entails *q,* then she should believe that *q*. It is the job of logic, statistics, probability theory, and decision theory to codify such constraints on thought. Thus, the interpretationist claim is that to interpret thinkers we must assume that they are consistent in these ways and we must take them to abide by such logical constraints. In this paper I shall only focus on logical consistency since all logicists demand this whatever other constraints they may recommend.[1]

One loose group of philosophers who adopt this approach are those who accept Quine's claim that all interpretation has the character of radical translation. This is the kind of interpretative exercise carried out by linguists who attempt to converse with

[1] There is some debate over which principles of logic and probabilistic reasoning are essential for rationality. Harman (1999), for example, argues that the principle of deductive closure is not constitutive of rationality. However, whether logical consistency is necessary for rationality is generally not contended: of course a rational thinker should not believe both that *p* and that not-*p*.

tribes or peoples with whom they have no previous linguistic connection. There are no etymological links — no common Graeco-Latin roots for example — that can be used as shortcuts in determining what they are talking and thinking about. There are no previous translations of this community or any obvious parallels in syntax or meaning between their language and others that have been successfully translated. All such linguists have to go on, therefore, is the current behaviour of the natives (or 'aliens'). Quine argues that for such interpretation to proceed we must accept certain principles of charity. Crucially, we must assume that our subjects' thoughts are logically consistent.

> Wanton translation can make natives sound as queer as one pleases. Better translation imposes our logic upon them…That fair translation preserves logical laws is implicit in practice even where, to speak paradoxically, no foreign language is involved. Thus when to our querying of an English sentence an English speaker answers 'Yes and no', we assume that the queried sentence is meant differently in the affirmation and negation; this rather than that he would be so silly as to affirm and deny the same thing….The common sense behind [this] is that one's interlocutor's silliness, beyond a certain point, is less likely than bad translation[.] (Quine, 1960: 58-9)[2]

Davidson also embraces radical translation and Quine's logicist assumptions.

> The process of devising a theory of truth for an unknown native tongue might in crude outline go as follows. First we look for the best way to fit *our* logic, to the extent required to get a theory satisfying Convention T, on to the new language; this may mean reading the logical structure of first-order quantification theory (plus identity) *into* the language, not taking the logical constants one by one, but treating this much of logic as a grid to be fitted on to the language in one fell swoop. (Davidson, 1973:136; my italics)
>
> [A]ll thinking creatures subscribe to *my* standards or norms of rationality. (Davidson, 1985: 195)[3]

[2] I shall argue below that thinkers can intelligibly be seen as believing contradictions. Here Quine considers whether we could ever translate the native as explicitly assenting to '*p* and not-*p*' (what Davidson calls an 'obvious contradiction'). This, however, is not a linguistic practice that I shall ascribe to intelligible thinkers. I shall only argue that we can interpret thinkers as having widespread contradictions in their beliefs of the kind where they believe that *p* and they also believe that not-*p*, and it would seem that Quine would also find such interpretations 'silly'.

[3] See also Davidson (1997:218): 'many philosophers have a hard time grasping why irrationality creates a conceptual difficulty; they regard someone who emphasises the tie between rationality and explicability, and the centrality of consistency in rationality, as an obsessed rationalist who cannot understand any form of reason not based on simple logic. I want to plead guilty, and throw myself on the mercy of my (largely rational) readers.'

In what follows I shall mainly focus on Davidson's account of interpretation rather than on Quine's since Quine is an eliminativist about belief, intentional content, and meaning.

Logicists take epistemic rationality to be codifiable in terms of certain logical constraints by which all rational thinkers must abide. This combined with interpretationism entails a certain constraint on intelligibility: the logicist insists that in order to be able to understand someone — in order to be able to make their thoughts intelligible — we must project *our* logic onto their thoughts; we must therefore strive to see them as not holding contradictory beliefs. Bortolotti (2005a) calls this the rationality constraint, and it is a constraint that I shall reject.[4] First, though, in the next section I shall consider the empathetic aspect of interpretation, that which is inherent in the claim that we must project *our* logic onto alien or native speakers.

3. Interpretation and Empathy

In order to be able to interpret the thoughts of the native or alien, I must be able to think like them. I do not have to share their beliefs, but I do have to be able to grasp the content of their thoughts, that is, I have to be able to understand — yet not necessarily agree with — the way that they think about the world.

> [When interpreting a thinker] we project ourselves into what, from his remarks and other indications, we imagine the speaker's state of mind to have been, and then we say what, in our language, is natural and relevant for us in the state thus feigned....what is involved is evaluation, relative to specific purposes, of an essentially dramatic act. (Quine, 1960: 218-9)

> You appreciate the reasonableness of an action by putting yourself into its agent's shoes, not by forcing him into yours. (McGinn, 1977: 522)

As the native utters 'gavagai' we consider what we would have said *if we were him*. We may perhaps have been drawn to utter 'rabbit' and so we forward this as a translation of

[4] Others have also argued against the rationality constraint. See, for example, Stich (1981), Cherniak (1986) and Bortolotti (2005b). Stich and Cherniak, though, do not take account of Davidson's claims concerning background rationality; I shall discuss these in section 7 below. It is primarily this aspect of his account that my argument aims to undercut. Bortolotti, however, focuses on Davidson's thoughts concerning recovery and these will also be discussed below.

his utterance. To check whether this is a good translation we can try out 'gavagai' in other situations in which we take 'rabbit' to be appropriate. If signs of approval are elicited from the native then we will make a start on our translation manual, if not then 'rabbit' would be rejected and further empathetic acts would have to be attempted.

Following McCulloch (1999) I will clarify just why this process should be seen as empathetic. Let us consider the home case in which we come to understand an utterance made by someone of our own linguistic community. As we interpret the words of a friend we do not just ascertain that he makes a certain noise: 'ra-bit'. Rather, we interpret him as saying that *a rabbit is over there*. The act of interpretation is more than an acoustic exercise; we must understand what our friend *means* by 'rabbit'. Thus, in order to interpret the thoughts of a thinker we must understand the content of the propositional attitude ascriptions that we make. And: 'Only if I can understand your words in your way can I gain full-blooded understanding of you as a subject of propositional attitudes.' (McCulloch, 1999: 34) Similarly with natives or aliens: to ascribe propositional attitudes to them we must come to understand 'gavagai' in the way that they do. And when we have, then we can go about finding expression to this in our own idiom. One understands and then one translates. Interpretation is thus an empathetic exercise since we must take on the alien's way of understanding his words and not simply use our own.[5, 6] I am in agreement with this claim concerning empathy; I will argue, however, that one can empathise with the alien — and therefore come to see him as intelligible — without being tied to logicism.

Such 'dramatic acts' of interpretation are also applicable to the thoughts of imaginary communities. In the next section my argument will involve the drawing up of an alternative community with which we can — in our imagination — come to share thoughts. Such thinkers are therefore intelligible. We shall see, though, that such a

[5] The continental tradition describes this empathetic aspect of understanding others in terms of hermeneutics or *verstehen*.

[6] Collingwood's (1939) thoughts on history also involve the claim that empathy plays a key role in interpretation: 'When I understand what Nelson meant by saying "in honour I won them, in honour I will die with them", what I am doing is to think myself into the position of being all covered with decorations and exposed at short range to the musketeers in the enemy's tops, and being advised to make myself a less conspicuous target...Understanding the words means thinking for myself what Nelson thought when he spoke them...Unless I were capable — perhaps only transiently — of thinking that for myself, Nelson's words would remain meaningless to me.' (1939:112)

community cannot be interpreted along logicist lines, and this is because the holding of contradictory beliefs is too widespread in this community to allow interpretation driven by a logicist construal of the principles of charity.[7] In order, though, to make such an alternative form of thought intelligible, I must provide a plausible background that explains why such intelligibility can be sustained.

4. Gullible

The inhabitants of City State Gullible — the Gulliblians — believe whatever they are told.[8] Knowing this, tourists are generally helpful and pass on truths about the world. However, in recent times tourists have become rather mischievous and for fun they feed false information to the Gulliblians. From their own empirical investigations the Gulliblians have formed the belief that the world is spherical, yet via misleading testimony they have also formed the belief that it is flat. They believe that *p and* they believe that not-*p*, and such contradictory beliefs have become widespread. This is the key feature of this community, and the coherence of this scenario is crucial to my argument against the logicist. I shall therefore spend most of this section showing just how this community can be seen as intelligible. First, it must be clear that we really do know what to say about the thoughts of such people; that their behaviour, although strange, is truly indicative of particular sets of contradictory beliefs. To confirm that this is so, one must remember that beliefs are not only expressed in explicit statements such as 'I believe that *p*', but that they are also manifest in the actions we perform, and it is through looking at the wider behavioural profile of the Gulliblians that I can substantiate my interpretation.

It has been noted in Gulliblian history that since the arrival of the tourists a certain indecision has become manifest. Sometimes, in cases where contradictory beliefs

[7] The holding of contradictory beliefs is seen by some philosophers as definitive of self-deception. Mele calls this 'dual-belief' self-deception as opposed to 'garden variety' self-deception in which a thinker falsely believes something that she would like to be true even though there is evidence that her belief is false. See Mele (2001) for a thorough treatment of both forms of self-deception and for further references to empirical and conceptual work on these phenomena. Here I shall not be concerned with self-deception itself, although, as suggested in section 5, evidence for such dual-belief cases would lend some support to my claims.

[8] Gulliblians exhibit what Goldman (1999: 109) calls 'BLIND TRUST' and what Fricker (1994:140) calls 'simple trust'.

are held, citizens cannot bring themselves to act. They may believe that it is Christmas and that it is not. They think that they should deck the halls with boughs of holly *and* that it is just the wrong time of year to do so. Thus, they spend more time than before standing around in a state of indecision. Some of their action has been eroded by the newcomers. However, if such indecision becomes too rife, then all attempts at interpretation may be frustrated: providing a determinate interpretation of their beliefs would be problematic since the behavioural manifestation of either the belief that *p* or the belief that not-*p* would be thwarted by the simultaneous holding of its contrary. It would be helpful, then, if there were more revealing evidence for the holding of particular sets of contradictory beliefs. And this there is: Gulliblians claim to expect no rain yet take their raincoats with them; they vehemently argue that meat is murder yet heartily tuck into steak tartare when dining with the tourists. They themselves can back up their professed beliefs with reasons, and, as we shall see below, we can support our additional interpretation of beliefs to the contrary with behavioural evidence in the context of what we have learnt about their moral code.

I am not claiming that such thinkers have reason to believe that '*p* and not-*p*', or that they could manifest a certain kind of behaviour that points towards the belief that '*p* and not-*p*'. Gulliblians merely have both the belief that *p* and the belief that not-*p*. It will therefore be different aspects of their behaviour that point to these respective beliefs and, at any one time, they may only be aware of at most one of them. On a Cartesian account of the mind, beliefs are episodic and they are essentially open to introspection. According to such an account the Gulliblians would have to be interpreted as believing *p* one minute and believing not-*p* the next, and as never having both beliefs at the same time. If, however, one accepts a dispositional account of belief then it can be allowed that a thinker can believe that *p* without being currently aware that he does so. A thinker can therefore be aware of believing that *p* while also unconsciously believing that *q*;[9] and, I claim, this is also the case with respect to the belief that *p* and the belief that not-*p*. Gulliblians need not be interpreted as believing *p* one minute and not-*p* the next; they

[9] Such talk of unconscious beliefs does not commit me to any particular theory concerning the structure of consciousness (to perhaps, say, a Freudian one); the claim is just that a thinker can possess beliefs of which he is not at the moment aware, e.g. you believe that Rome is in Italy although you were not conscious of this belief a few seconds ago.

can be seen as simultaneously having both beliefs but, at any one time, as only being aware of at most one of them. When one is aware of a certain belief then we can say that it is operative or salient. Given such an account we should now think again about the Gulliblian beliefs concerning Christmas. First, coherent action is not necessarily thwarted because at a certain time a Gulliblian can act on whichever belief happens to be salient. And second, when action is thwarted, this is not because there is no action appropriate given the possession of the belief that '*p* and not-*p*'; rather, their indecision is caused by first one belief becoming salient, and then the other. (And such oscillation is something with which we can easily emphasise: she loves me; she loves me not…)

You may initially find such a scenario incomprehensible: we do not have intelligible thinkers here and the best we can do is offer a pathology of how and why logical thought has broken down. I, however, shall offer more and I shall argue that these thinkers are intelligible. Before substantiating my claim I should make it clear just what I mean by this. One can give an explanation of behaviour by describing how it is caused, and sometimes one can also show that such behaviour is intelligible by citing reasons for why a certain course of action was pursued. These are distinct explanatory strategies. The strategy with which I am concerned is the latter. Following the former strategy potentially any behaviour can be explained, be it irrational, seemingly random, or alien to us in some way. Clinical psychologists and biologists can give such explanations of the behaviour of the psychopath and the termite. To give a reason-backed explanation, however, one needs to grasp one's subjects' reasons for acting, and there are limits to one's empathetic ability to do this, limits that the psychopath and the termite transgress. The community of Gullible, however, is different: we *can* come to understand the way that they think by grasping their reasons for acting and for saying what they do. Their behaviour is intelligible to us as well as causally explicable.

As argued, understanding essentially involves empathy: I can only see you as an intelligible thinker if I can come to imagine myself having your thoughts. So too with the alien community: we can only come to see them as intelligible if we can imagine ourselves thinking like them. And in order to help us to do this I will sketch some of the moral principles of Gullible. The moral backdrop that I present will encourage the holding of contradictory beliefs.

In our community young children manifest blind trust in the testimony of others. As we get older, though, we lose some of our openness to testimony as we discover error and deceit. This, however, is not so in Gullible. To the Gulliblians trust is highly prized. It has been raised to be the paradigm virtue by which they should live. Their ancient philosophers found man to be a 'trusting animal' and ever since there has been a cultural aspiration to live by that ideal. Not to believe someone's word is seen as excessively discourteous and uncivil, and those who do not trust others are expelled from the community.[10] A later philosopher, Gant, argues that the core of morality lies in the treatment of individuals as ends in themselves, and that proper respect for a person is only shown if one does not lie to them and if one uncritically believes what they have to say. Distrust would reveal that one's interest is simply in the truth of their utterances and that the utterer is merely a means for you to acquire that truth. Trust, therefore, is seen as more important than logical consistency. Citizens blindly trust both the tourists and their friends within the community. Inevitably, then, they are forced to take on contradictory beliefs and this they must do if they are to live a moral life.

An analogy can be drawn here between Gullible and a community of radical chess players. In chess, good moves are those that develop one's position in order that the opponent's king comes to be threatened either through strategic or material advantage. Such moves can be seen as analogous to the performance of rationally justified linguistic acts; in both we are aiming at a certain standard of correct play. *We* see consistency in our chess moves as essential to such play: the rules of movement and capture must be strictly adhered to. There is a logic to chess play as there is to language. Imagine, though, a radical group of players: they allow considerations other than consistency to direct their play; not moral considerations, as with the Gulliblians, but ones that are aesthetic. Moves can break with the rules if the positions to which they lead are aesthetically pleasing in some way: there may be an elegance or a certain humour in the lines produced. Not anything goes, however; the rules of movement and capture can only be broken with regards to aesthetic considerations. And, importantly,

[10] See Shapin (1994). He discusses the class and role of 'gentlemen' in early 17th century English science. In such a society a gentleman's word was his bond and to query this or to distrust him was to doubt not only his word but the fact that he was a gentleman. 'To distrust, or not believe, is to Dishonour.' (Hobbes, *Leviathan*, 10.27)

the game does not simply transpose into art; the aim is still to kill the King. For the radicals, then, logical consistency is no longer their sole guiding constraint.

It may be objected that the radicals are just not playing chess since chess is constituted by the set of rules that govern it, rules that are blatantly broken by these players. There is, however, a variant of my scenario in which the rules of chess are not contravened. The radicals may introduce aesthetic considerations within the context of our current rules. The rules of movement and capture may not be broken, but plodding or even strategically poor moves (by *our* standards) are seen as good moves if they lead to elegant, beautiful or funny positions. The radical game is adopted and grand masters are those whose primary concern is with aesthetics and not material or strategic advantage. The analogy with Gullible is that for the Gulliblians logical consistency can also be overridden; for them, however, it is overridden by moral rather than aesthetic concerns.[11]

The Gulliblian's assertions concerning vegetarianism and the ethical arguments forwarded to support his claims can be taken as evidence for his belief that eating meat is wrong. In ascribing this belief to him I do not have to consider whether the contradictory belief is held. Further beliefs are also revealed in the French restaurant. His unconcerned and greedy consumption of the bloody steak indicates that he has the belief that eating meat is not wrong. And similarly in ascribing this belief I do not have to consider whether the contradictory belief is held. Further, through considering what it would be intelligible for us to think if we lived in their society, we can come to interpret and explain why this thinker has both beliefs and how he manages to shield these two beliefs from each other.

5. Moral, Psychological, and Pragmatic Considerations

As well as the game of radical chess, there are three further lines of support for the possibility of a community such as Gullible. These involve broadly moral,

[11] See also Michel Platini on Eric Cantona (football players): 'One gets the feeling that if he can't score a beautiful goal he'd rather not score at all' (attributed). And when Cantona was asked whether he is interested in going into management he replied: 'If I do it, I want to see a team play football as it has never been played before. I want it to be like an artist with a new movement, like a revolution.' (*The Guardian*, Saturday, June 23, 2001, pp. 4-5). It seems, then, that he is indeed a radical.

psychological, and pragmatic considerations.

First, it is easy to empathise with a moral constraint on trust because this is something that acts on us and something that has had, in various cultures and times, an even stronger hold over citizens.[12] I do not trust what you say just because doing so may be pragmatically advantageous to me; there is, rather, a moral dimension to trust: trusting you is a way of treating you as a person. Trust in others is more than mere reliance: I may trust the rope to hold my weight, and I trust you to take me safely to the top of the mountain.[13] If the rope breaks I am annoyed (or worse); if, however, you abandon me, I am not just annoyed with you — I would feel that you have to some extent forsaken your humanity. I would be annoyed *and* I would disrespect you.[14] And we should not be distracted here by the high stakes example: prima facie I would feel disrespect towards you if you failed to turn up at the coffee shop after you said that you would ('prima facie' because there may be a good explanation for you not doing so). Conversely, if someone doesn't accept what we say, then in many circumstances we would feel slighted and think that person rude. The moral dimension of trust is also evident if I consider how I would react if questioned about why I didn't accept someone's word for something: in certain cases I would feel embarrassed. In any situation there may of course be a wealth of reasons why I do not accept what someone says: I may know more about particle physics than them, or I may have seen the latest weather forecast. However, the having of such reasons alone does not excuse me from embarrassment and moral censure; I would also need to explain to them why their word was not being taken on trust. Testimony, then, has a moral dimension and one to which we are all sensitive.

A stronger claim is that this moral attitude of trust is compulsory. We naturally trust others and feel disrespect towards those in whom our trust turns out to be

[12] See Montaigne: 'Our minds never work except on trust, they are bound and controlled by their appetite for another man's ideas, enslaved and captivated by the authority of his teaching'. (1958: 55)

[13] Such trust in the rope is based on its empirical reliability: I trust it because I know that ropes made by Mammut have been strong in the past. Trust in others, however, does not seem to be so tightly indexed to reliability. Don't you sometimes trust a friend even though he may have regularly let you down in the past? (If, however, he lets you down enough he may cease to be your friend and then he doesn't demand your trust. There is, though, an interesting amount of slack here: there's a period where you should still trust him even though he is not reliable.)

[14] Holton (1994) is good on the distinction between reliance and trust.

misplaced. Trust can be seen as what Strawson calls a reactive attitude.[15] These are attitudes such as resentment, praise, gratitude, and love that we naturally hold towards our fellows and it is the possession of these that is constitutive of being a person. Someone who does not feel them may be a member of our biological species, but they are not a person. The claim, then, is that trust is a necessary part of personhood.[16]

Here we have been discussing our own society and not Gullible; Gullible, therefore, is just a society in which this moral constraint is more pressing and thinking of their society in this way may help us to imagine living there and thinking like them.

Second, there is psychological evidence that people can have contradictory beliefs, and such contradictions are part of both folk psychology and scientific psychology.[17] A physics student may spend his day calculating that a motorcycle (a mass of so many kilogrammes) cannot make a corner of a certain angle when moving at a certain velocity. However, when motorcycling home from work he continues to ride into such a corner at such a velocity on a bike of that mass. He therefore believes both that *p* (that his bike will make the corner) and that not-*p* (that it won't). He only though focuses on one of these beliefs at any one time and therefore has that characteristic oscillation between panic and calm. Such self-deception is facilitated by our ability to ignore or misinterpret certain evidence for not-*p* and selectively to focus on evidence for *p*. When thinking of his experience of similar corners in the past he puts his calculations to the back of his mind; he highlights the positive evidence and hides the negative evidence.[18, 19]

It should be noted, though, that all such attributions of self-deception are contested. Mele (2001) provides alternative interpretations of the kinds of examples

---

[15] See Strawson (1962) and Holton (1994).

[16] We can sometimes suspend our reactive attitudes and take the objective attitude towards others. In doing so we would only rely on others acting in a certain way, perhaps in the way that they always have, or in the way that they say they will; we would not, though, *trust* them in the moral sense we have been discussing.

[17] Here I have concentrated on the holding of contradictory beliefs, but there is also compelling evidence that thinkers do not abide by other aspects of the logicist conception of rationality such as deductive closure and certain axioms of probability theory. See, for example, Tversky and Kahneman (1981) and Cherniak (1986).

[18] For various other plausible examples of self-deception, see Werth and Flaherty (1986).

[19] He does not do this because he wants to believe that he will make the bend. If he did there might be problems here with the notion of willing to believe (see Williams, 1973). Rather, he ignores the

cited in the literature. He does not, though, have any a priori reasons for doubting the possibility of self-deception; he only argues that there are more plausible alternative empirical interpretations of such cases.[20] In my imaginary case of Gullible, however, the empirical evidence in favour of self-deception is bolstered. Even if Mele's interpretation of real world cases is correct, the possibility of Gulliblian self-deception is still open because of the role that trust plays in their society.[21]

We must be careful to note a crucial difference between the Gulliblian scenario and that of the physicist biker. The latter is a member of a community who in general do not have contradictory beliefs. His inconsistent beliefs are thrown into relief against a background of consistency. Thus, as we shall see below, Davidson and the logicists would claim that to interpret him we must work with the assumption that he does not hold contradictory beliefs, and only by doing so can we begin to build up an overall coherent interpretation of his thinking. Once such a background is established, however, anomalous episodes of behaviour may suggest that we should overrule our working assumption in certain cases. The fact that the biker maintains his speed going into the corner suggests that we should overrule our assumption in this case and allow that he believes both that *p* (that he'll make the corner) and that not-*p* (that he won't). In Gullible, however, we will not get very far if we work with the default assumption that evidence for the belief that *p* is also evidence that a thinker does not believe that not-*p* (and vice versa). If we cannot ascribe a certain belief where there is evidence for the holding of the contrary belief, then the widespread self-deception in Gullible entails that most of their thoughts are impenetrable to us, and thus we will not be able to build up a background interpretation against which episodes of self-deception become manifest. An alternative interpretative stance must therefore be adopted; with them the logicist constraint that one cannot intelligibly believe both that *p* and that not-*p* must be

---

conclusions of such calculations since they cause him anxiety, and such anxiety-reducing measures may come naturally to him. See Mele (2001:58-9).

[20] 'I have no wish to claim that it is impossible for an agent to believe that *p* while also believing that not-*p*. My claim is that there is no explanatory need to postulate such beliefs either in familiar cases of self-deception or in the alleged cases cited by these researchers and that plausible alternative explanations of the data may be generated by appealing to mechanisms and processes that are relatively well understood.' (Mele, 2001: 92)

[21] Orwell also provides fictional support for the possibility of self-deception in his *Nineteen Eighty-Four*. He calls such self-deception 'doublethink'; it is the holding 'simultaneously two opinions which cancelled out, knowing them to be contradictory and believing in both of them.' (1984: 186)

dropped.

Third, there are pragmatic considerations that support the possibility of Gullible. William James argues that too much concern with whether our beliefs are true could lead to an overly cautious strategy of belief acquisition and consequently to an impoverished set of beliefs; it may be better to take epistemic risks in order to perhaps ultimately arrive at a wider set of truths. '[T]he risk of being in error is a very small matter when compared with the blessing of real knowledge.' (James, 1948: 100) Rescher and Brandom (1979) also argue for such a pragmatist approach. They, however, specifically mention the holding of contradictory beliefs. It may be better to have a wider body of knowledge that includes some inconsistent beliefs than to be cautious and to only have a narrow set of beliefs which one is sure does not harbour any contradictions.[22]

These pragmatists focus on other epistemic goals such as error-avoidance and breadth of knowledge (those, that is, that are ultimately truth related). My claim, however, is that the goals that allow for contradiction within one's belief system need not be directly epistemic; they can be moral. Philosophers have a tendency to be rather precious with respect to truth; they shouldn't be — truth isn't all that we seek and it is not the only goal that shapes our belief systems. Here is Plantinga questioning this focus on epistemic goals.

> Questions arise here: first, do we *all* have this goal? Might not some of us have no overarching cognitive goals at all? Might not others have different overarching cognitive goals — comfort, say, or salvation, or fame or fortune, or mental health? Might I not have the primary goal of doing my cognitive duty, *hoping* that this will lead to my holding mostly true beliefs, but not explicitly aiming at the latter? (Plantinga, 1993: 97)

I accept what could be called a Humean account of intelligibility: intelligible action and thought is that which can be interpreted as aimed at the satisfaction of one's desires. Such intelligibility will usually be manifest in a subject's beliefs not being contradictory

[22] See also Simon's (1957) notion of 'satisfice', and Hookway (2006:100): 'although deductive logic tells us about the relations between the truth values of different propositions and must inform our alethic judgements, logic does not tell us what inferences to draw and what beliefs to form. The latter reflects which true propositions are worth attending to, which of them are likely to be relevant to our other concerns, and how ready we are to risk conclusions that will subsequently turn out to be incorrect.'

since '[t]he man who believes both that it is raining and that it is not is badly placed to act as if he wants, say, to avoid getting wet. (Blackburn, 1993: 190) However, given the right circumstances contradictory beliefs may be required in order that one's goals and desires are satisfied. If not getting wet is all that concerns the Gulliblian then it would be advantageous to him not to have contradictory beliefs concerning the weather; but the Gulliblian is always primarily concerned with treating his fellows with respect and thus trusting whatever they say. If these are his driving desires then we should interpret him as holding contradictory beliefs. Logic, therefore, does not necessarily override all the other social, causal, and pragmatic constraints and pressures on belief.[23]

6. Are Gulliblians Believers?

I have suggested that it is an empirical question whether thinkers could hold contradictory beliefs and Gullible has been forwarded to show that such beliefs are possible. There are, though, various arguments against the coherence of the Gulliblian scenario, arguments that attempt to show that we are losing sight of the very notion of what it is to have a belief. I shall look at four such arguments, all of which I reject.

First, it could be claimed that Gulliblians do not *believe* what the tourists tell them because they have a weaker attitude towards what they say. Some philosophers suggest that 'acceptance' may be weaker than belief in terms of the degree of conviction one has towards that attitude; it is, however, strong enough to guide action.[24] I may, for example, accept that spinach is good for me – and thus regularly eat it – yet not believe that it is. I could accept that a free market economy is the only plausible way to run the modern world, yet not believe that this is so. Gulliblians could therefore accept what the tourists say without being guilty of holding contradictory beliefs.[25] I can, though,

[23] A logic professor may find all kinds of logical mistakes made by his friends and students silly, but he should nevertheless interpret his fellows as making such mistakes given what he knows about their thinking and the prevalence of such errors in non-professionals (see Cherniak, 1986: 96-7). Quine's 'translation policy' on p. 4 above should therefore be driven by the claim that 'better translation does not attribute acceptance of inconsistencies that are obvious *for the subject*'. If one accepts this, then the global logical errors in Gullible are interpretable and intelligible given what we know about their moral code and psychology.

[24] Various philosophers mark a distinction between acceptance and belief. See Cohen (1992), Velleman (2000: 250-255), and Bratman (1999:15-34).

[25] One such account of belief and acceptance is given in Cohen (1992). His account, though, is not conducive to this criticism of the Gullible scenario. Cohen argues that belief amounts to a disposition to

stipulate that this distinction does not fit the psychological facts in Gullible. The Gulliblians can have whatever level of commitment is required for belief, and they can have it for both their belief that *p* and their belief that not-*p*. They may not, however, be conscious of their strong commitments to both *p* and not-*p* at the same time because their awareness of such commitments can fluctuate over time as one belief and then the other becomes salient.

Second, Davidson claims that it is a conceptual truth that a belief with a certain propositional content bears certain specific logical relations with one's other mental states and beliefs. This is an a priori claim and since it is *our* concept of belief we are talking about, then it is *our* laws of logic that are conceptually necessary for the existence of such mental states. The Gullible scenario is therefore incoherent because it is conceptually impossible to have the belief that *p* if one also has the belief that not-*p* (and vice versa).[26]

My response is to accept that one's beliefs and other propositional attitudes must be logically related together in some way and that they must have certain causal links to the world. I deny, however, that there is a necessary set of such relations. In order to see an alien as a believer we have to be able to empathise with his thinking; we have to come to think like he does and so there has to be enough in common between our thinking and his to enable us to do this. And this there is: a story has been told that enables us to empathise with Gulliblian thought. There is much that our thinking does have in common with that of the Gulliblians: their beliefs are causally sensitive to perceptual evidence, to what they learn from others, and to certain moral pressures they face; and their beliefs guide their actions. It is the act of empathising with such aspects of their thought that enables us to be able to interpret them, and the limits on what we

---

feel that *p*, and thus on his account it is OK to hold both the belief that *p* and the belief that not-*p* (1992: 31-3) since one's disposition to feel that *p* can be instantiated on different occasions to one's disposition to feel that not-*p*.

[26] See also Dennett (1987). He argues that one cannot ascribe contradictory beliefs by adopting the intentional stance and therefore that the possession of such beliefs does not make sense. The behaviour of such thinkers would have to be explained using the design or physical stance and any such explanation would not involve *belief*: "[W]e must descend from the level of beliefs and desires to some other level of theory to describe this mistake, since no account in terms of beliefs and desires will make sense completely. At some point our account will have to cope with the sheer senselessness of the transition in error…This is not to say that we are always rational, but that when we are not, the cases defy description in ordinary terms of belief and desire." (1987: 87)

can empathise with are not delineated by a set of logical constraints that underpin *our* language.[27] There is not a 'common core' or fixed 'bridgehead' consisting of such constraints (Hollis, 1982: 74). There is, instead, a floating bridgehead or pontoon since there must be enough in common between their thoughts and ours to allow interpretation to proceed. The claim that there is this much constraint on interpretation is a priori; the nature of this pontoon, however, is an empirical issue and the interpretative theory that is most successful at explaining and understanding (and perhaps predicting) native or alien behaviour is the theory that should be adopted.

Third, it could be suggested that we cannot coherently see the Gulliblians as believers because their thinking is not responsive enough to the world; their attitude to the testimony of others is too resistant to revision and so does not therefore amount to belief. This, however, ignores the common sense picture of beliefs which allows that one can stubbornly hold onto them in the light of contradictory evidence. I still believe that Alex Higgins will once again win the world snooker championship even though by any reasonable lights I should have given up this belief long ago.[28] And moon landing conspiracy theorists still hang onto their beliefs against all the evidence. Such beliefs can be seen as unreasonable or irrational but they are still beliefs. If anything Gulliblian thinking is *too* responsive to the world: they form beliefs based on their own empirical evidence *and* they are highly sensitive to testimonial evidence whether or not this is consistent with the beliefs that they have acquired as a result of their own investigations.

And fourth, it could be claimed that the Gulliblians have lost sight of the essential normative character of belief. Beliefs are mental states with representational content; they are aimed at the truth. There are therefore certain beliefs we *should* have or those we *ought* to have given that we aim to represent the world correctly. And again this cannot be so given the manifest contradictions that are present in Gulliblian

[27] It is not therefore the case that we must find the alien 'consistent, a believer of truths, and a lover of the good (all by our own lights, it goes without saying)' (Davidson, 1970: 222). Davidson's principle of charity should be rejected. Others have also argued to this conclusion, and Grandy (1973) suggests replacing Davidson's principle with a principle of humanity. We need not find truth and consistency in the thoughts of others, but we do need to interpret others as having thoughts as similar to our own as possible. I am more sympathetic to this constraint although it is important that it incorporates both empathetic *and* imaginative elements. Gulliblian thought is in certain ways not very similar to our own — I may not trust my neighbour — but our thinking would be similar *if* we had a moral outlook like theirs.

[28] For empirical evidence concerning such 'belief-perseverance' see Ross, Lepper and Hubbard (1975), and Ross and Anderson (1982).

thought. The Gulliblians may have an ethical commitment to believe the tourists but they do not have the required epistemic commitment since in many cases they should not believe the tourists because of the overwhelming evidence against what they say.

I shall sketch two broad responses to such an objection. First, it could be maintained that the Gulliblians do have a normative commitment and that they still chase the goal of true belief. Their eager acceptance of testimony not only helps them manifest the trust they value so highly but, all being well, it helps them acquire truths about the world. The new tourists may thwart this goal, but that is no fault of the Gulliblians; their aim is still on the truth. It is consistent to claim both that beliefs are aimed at the truth and that sometimes they can miss the target. And in the case of Gullible, testimonial beliefs can systematically miss because of the diverting influence of trust.

Second, a more radical response is to claim that normativity is not essential to belief. Beliefs can be true and they can be false, but a believer need not always strive to have true beliefs.[29] The claim, then, would be that belief ascription is purely descriptive and not normative. Here I follow Åsa Wikforss (2001: 205). With respect to the alleged normativity of meaning she says:

> That an expression is true of some things but not of others does not in itself imply that it should be used in any particular way, that there are 'normative truths' about my behaviour with that expression. If 'horse' means *horse*, then 'horse' is true of horses only, but it does not follow that I *should* apply the term to horses only.

The Gulliblians' commitment to their beliefs is driven partly by moral trust and not wholly by truth. Gulliblians may therefore have more false beliefs than us but they are still believers since it is not the case that believers *ought* to believe the truth. Beliefs are just representations of the world — some true; some false — but it is not the case that we ought to have the former and not the latter, even though we quite often do. And note that this is consistent with Davidson's logicist account of interpretation and intelligibility. The content of a thinker's beliefs is determined by what he is seen to hold

[29] That is not to deny that when one believes that *p* one believes that *p* is true. One can accept that this is a conceptual truth while also claiming that thinkers are not under a normative obligation to believe that *p* when *p* is true.

as true, and Davidson does not argue that a thinker *should* hold certain claims to be true, only that he does and his doing so enables us to acquire an interpretation of his beliefs and desires. The extra, and unnecessary, logicist strand in this picture is that the holding true of contradictory claims thwarts interpretation. I can deny this without rejecting other aspects of Davidson's interpretationist approach.

I have considered four ways to reject the claim that Gulliblians have contradictory beliefs. None of them, however, are persuasive. I have argued that Gulliblians are believers; believers whose beliefs bear certain logical relations with their other mental states, and are responsive to the world. Some of their beliefs are admittedly resistant to revision and perhaps devoid of normative content, but this does not impugn their very nature as beliefs.

7. Davidson, Epistemic Crisis and Recovery

Davidson does accept that thinkers can sometimes be interpreted as having contradictory beliefs. I shall now sketch how he can allow this; how his account is still driven by logicist assumptions; and how Gullible remains a problem.

Let us first remind ourselves of why contradictory beliefs are problematic for Davidson. The content of a mental state is that which it can be interpreted as having in the context of our interactions with the world and each other. And a belief that *p* is the kind of intentional state that can be ascribed along with the belief that *p* entails *q*, and the belief that *q*. To have a belief that *p* is simply to be interpretable as having a state that lies in such logical relations with one's other intentional states. However, the holding of contradictory beliefs seems to threaten the practice of interpretation, the practice that, according to Davidson, is the only handle we have on the identity of intentional states. If one also possesses the belief that not-*p* then it is not clear how the belief that *p* can be unambiguously manifest in one's behaviour.

Thinkers can however be interpreted as holding contradictory beliefs if three conditions are satisfied. First, the majority of a thinker's beliefs must be non-contradictory. This allows an interpreter with his logicist assumptions to build up a consistent background picture of the thinker's thoughts. Once this has been done local deviations from consistency can be identified.

It is still the case, though, that a thinker must not be aware of any such contradictions. Second, therefore, there must be mechanisms that allow contradictory beliefs to be hidden or shielded from each other. This could be achieved if the mind were divided or partitioned in some way.[30] Such partitioning could be seen as involving some kind of Freudian or Platonic apparatus, although less metaphysically problematic abilities would also suffice. All that's required is that a thinker must not be aware of both of his contradictory beliefs at the same time, and for this to be so he would only require some kind of divided attention or mechanisms that enabled him to reflect on his belief that *p* without his belief that not-*p* 'coming to the surface'.

Third, a thinker must be able to recover from his inherently unstable cognitive state, that of holding contradictory beliefs.[31] It may happen that the mischievous tourists cause the Gulliblians to fall into something of an 'epistemological crisis'.[32] As said, at times their contradictory beliefs thwart any coherent action on their part and, even when they do act, much confusion can follow: 'hold on, why did I order a steak tartare, I'm a vegetarian'. Ongoing reflection upon the state of their own belief system will result in them being able to pursue less and less of their thoughts and actions. Self-reflection may reveal to the Gulliblians that they have contradictory beliefs and that coherent action has therefore become problematic. The key question, then, concerns what they should do about this. Crises can be overcome and their epistemologists may realise that their problems are due to the clash between blind faith and their belief in their own empirical discoveries. At least two paths are open to them. First, if we went to Gullible we may be able to persuade them that blind faith is misguided and that a more measured form of trust would be advisable. Second, if their moral obligations prove to be too strong, they could perhaps reject their own findings completely ('my eyes and ears must be deceiving me; it cannot be raining because the tourists have assured me that it is fine'). Both of these strategies are consistent with Davidson's claim concerning recovery and, for him, we must pursue one of them.

There is, however, a third reaction to epistemic crisis and that is acquiescence.

---

[30] See Davidson (1985) and (1997).
[31] See also Dennett (1987: 95): 'Inconsistency, when discovered, is of course to be eliminated one way or another.'
[32] See MacIntyre (1980).

Elaborate and widespread systems of psychological suppression and shielding could be developed that enable them to avoid coming face to face with their contradictory beliefs. Gulliblian thought could therefore be interpreted even if the Gulliblians do not satisfy Davidson's third constraint on the interpretation of contradictory thought, even if, that is, they do not recover from inconsistencies in their beliefs.[33]

I shall not linger on this response to Davidson because my main argument has concerned Davidson's first constraint, the claim that inconsistent thought can only be seen as intelligible against a background of consistency. He claims that the assumption of non-contradiction is essential if we are to begin the interpretation of a thinker or community. Once begun, though, interpretation can allow for some contradictions — contradictions, though, that thinkers must be keen and able to recover from. I deny both of these claims. Non-contradiction need not be assumed either at the start or during interpretation, and the ability and desire to recover are not necessary for coherent thought. Even though Davidson allows for the possibility of contradictory beliefs, his logicist attitude towards them plays a central role in his account of interpretation. Their essential role in the radical case — at the start of interpretation — and the assumed need for recovery show that non-contradiction plays a constitutive role in what it is to be a thinker. I have argued that this is not so.

Cherniak (1986) also argues against logicism and agrees with me that we do not have to assume that thinkers are completely consistent; for him, though, they must satisfy a 'minimal consistency condition': "If A has a particular belief-desire set, then if

[33] Bortolotti also finds fault with Davidson's claims concerning recovery (see Bortolotti 2003). The ability to recover is dispositional and it is not therefore manifest in one's behaviour *now*; Davidson, however, holds there to be a constitutive link between interpretation and mindedness, and those essential aspects of thought — i.e. the holding true of non-contradictory claims, or the holding true of contradictory claims *and* the ability to recover — must be manifest in overt behaviour that is assessable by an interpreter. Bortolotti claims that this is not so with respect to recovery. Davidson does not however have a problem here. He takes interpretation to be a practice that takes time and, if it is to play a constitutive role in an account of the mind, one that must be seen as idealized. The beliefs of a thinker are those that an ideal omniscient interpreter can be seen as uncovering over time. There may be independent problems with the notion of an omniscient interpreter but, ignoring these, the fact that recovery is dispositional is not itself a problem for Davidson. Bortolotti also argues against the rationality constraint using actual case studies of people with delusions (see Bortolotti 2005a). She refers to the case of a person who claims to be first a man and then a woman and who is not willing to give up this contradiction even when it is pointed out to him (or her!?). This person holds contradictory beliefs and has no propensity to recover. Bortolotti also suggests that such a subject believes an obvious contradiction (i.e. they believe that '*p* and not-*p*'); as said, though, my argument does not depend on any such claim.

any inconsistencies arise in the belief-set, A would sometimes eliminate some of them." (1986: 16)[34] This is of course very vague (would Cherniak be happy to call a thinker rational if she eliminated just one of her widespread inconsistencies?). More importantly, though, Cherniak's position is distinct from the one that I have argued for. First, I reject the minimal consistency condition and argue instead for a minimal intelligibility condition. Consistency can be discarded so long as other common features of thought are maintained; enough, that is, to enable us to provide a determinate interpretation of the thoughts of the inconsistent thinkers. Cherniak does not agree:

> if an agent's cognitive system was not subject to some consistency constraint, and so could contain an unlimited number of inconsistencies, the attribution of such a system could not be of any value in predicting the agent's behaviour. (1986: 16)

I have argued that this is not so. Second, Cherniak rejects logicism because he claims that given our finite capacities, and thus our inability to survey the whole of our belief system, we cannot be expected to discover and subsequently to recover from all of the inconsistencies in our beliefs. He accepts, though, that this is what an ideal thinker would do. I disagree: even if Gulliblians were not limited by their finite capacities, and if it were possible for them to recover from all inconsistencies, they needn't.

8. Conclusion

Even though Davidson stresses the holistic nature of interpretation, the logicist constraints that we must abide by are given — they are constraints on the thoughts of all intelligible thinkers. I, though, have suggested a more extensive holism. One must be prepared to reject *any* feature of a belief system if there are unusual enough circumstances that would allow for a successful interpretation of alien thought which lacks that feature.[35] Our interpretation of Gullible cannot get off the ground if we remain wedded to our usual conception of intelligibility. To interpret this alien community the norms of intelligibility we use must not include the proviso that one should not have

[34] And other minimal constraints such as rational thinkers must abide by deductive closure *some* of the time.

[35] See Quine (1953:43): 'no statement is immune to revision. Revision even of the logical law of the excluded middle has been proposed as a means of simplifying quantum mechanics.'

contradictory beliefs. Non-contradiction is not therefore a precondition of interpretation: sometimes we can only understand an alien culture if we reject the assumption that intelligible thinkers do not contradict themselves.

We must, however, be careful not to be too liberal in what we see as intelligible. After all, Gullible is an imaginary community and can't we imagine *anything*: a community, perhaps, that believes everything it's paid to believe; or one whose water supply is contaminated with a logic drug that affects their modus ponens reasoning: if they think that *p,* and that *p* entails *q,* then according to their canons of reasoning they should also believe that not-*q* (and not *q*).[36] But can I really see such thinkers as intelligible? I have claimed that empathy must play a key role with respect to this question. I accept the Quinean line that unless I can think like the natives and the aliens, I cannot come to see them as intelligible thinkers, even though I may be able to predict how they will act. And perhaps I simply cannot imagine coming to think in the ways just suggested.[37] We can, though, empathise with the way that the acquisition of Gulliblian belief is influenced by their moral code and thus Gulliblians *are* intelligible. We have been urged to empathise with the inhabitants of Gullible, and radical chess players, physicist bikers and trusting mountaineers have been suggested in order to help us do this.[38]

The logicist aims to describe the contours of *all* intelligible thought. I have argued, however, that logicism is in fact parochial in its outlook. It takes *our* folk psychology as its benchmark and from this draws up a logical description of this manifest realm, a description, which it is claimed, captures the essence of intelligibility. My line has been to encourage one to transcend this parochialism, to transcend what is in fact merely a local realisation of intelligible thought. Logical consistency — that which underpins our thinking — is a cognitive virtue, but there are others, and what we see as virtuous need

[36] Or a grue-drug that induces everyone to believe that all emeralds are not green but grue.

[37] A related question concerns whether or not we can empathise with or understand those who are evil. See Morton (2004) and Hume's *Enquiry,* II.15: 'A man of mild manners can form no idea of inveterate revenge or cruelty; nor can a selfish heart easily conceive the heights of friendship and generosity.'

[38] As noted at various points, Bortolotti's claims are in many ways consonant with mine. She does not, though, place enough emphasis on empathy. For her, alien thought is intelligible if it can be explained and rationalised in terms of belief and desire. I, however, accept the Quinean claim that beliefs and desires cannot be interpreted without the interpreter's empathetic engagement with such mental states, a claim that we have seen drives Davidson's interpretationism.

not be accepted by other communities.[39]

> [T]he philosopher [should] in particular be alert to deflate the pretensions of any form of enquiry to enshrine the essence of intelligibility as such, to possess the key to reality. (Winch, 1958: 102)

Daniel O'Brien

University of Birmingham

D.Obrien@bham.ac.uk

[39] Thanks to Iain Law, Joss Walker, Lisa Bortolotti, Alex Miller and the staff seminar at Birmingham.

Bibliography

Blackburn, S. (1993), *Essays in Quasi-Realism,* Oxford, Oxford University Press.

Bortolotti, L. (2003), 'Inconsistency and Interpretation', *Philosophical Explorations* VI (2): 109-123.

Bortolotti, L, (2005a), 'Delusions and the Background of Rationality', *Mind and Language*, 20 (2): 89-208.

Bortolotti, L. (2005b), 'Intentionality Without Rationality', *Proceedings of the Aristotelian Society* CV (3): 385-392.

Bratman, M. (1999), *Faces of Intention: Essays on Intention and Agency*, Cambridge University Press, Cambridge.

Cherniak, C. (1986), *Minimal Rationality*, Cambridge (Mass.), MIT Press.

Child, W. (1994), *Causality, Interpretation, and the Mind*, Oxford, Oxford University Press.

Cohen, J. (1992), *An Essay on Belief and Acceptance*, Oxford, Clarendon Press.

Collingwood, R.G.(1939), *An Autobiography*, Oxford, Oxford University Press.

Davidson, D. (1970), 'Mental Events', in his *Essays on Actions and Events,* Oxford, Oxford University Press, 1980.

Davidson, D. (1973), 'Radical Interpretation', in his *Inquiries into Truth and Interpretation,* Oxford, Oxford University Press, 1984.

Davidson, D. (1985), 'Incoherence and Irrationality' in his *Problems of Rationality*, Oxford, Oxford University Press, 2004.

Davidson, D. (1997), 'Who is Fooled?', in his *Problems of Rationality*, Oxford, Oxford University Press, 2004.

Dennett, D. (1987), *The Intentional Stance*, Cambridge (Mass.), MIT Press.

Fricker, E. (1994), 'Against Gullibility', in Matilal & Chakrabarti (eds.), *Knowing From Words*, Dordrecht, Kluwer Academic Publishers, 125-161.

Goldman, A. (1999), *Knowledge in a Social World*, Oxford, Oxford University Press.

Grandy, R. (1973), 'Reference, Meaning and Belief', *Journal of Philosophy* 70, (14):439-52.

Harman, G. (1999), *Reasoning, Mind and Meaning*, Oxford, Oxford University Press.

Hobbes, T. (1994), *Leviathan*, ed. E.Curley, Indianapolis, Hackett.

Hollis, M. (1982), 'The Social Destruction of Reality', in M. Hollis and S. Lukes (eds.), *Rationality and Relativism*, Cambridge (Mass.), MIT Press.

Holton, R. (1994), 'Deciding to Trust, Coming to Believe', *Australasian Journal of Philosophy*, 72 (1): 63-76.

Hookway, C. (2006), 'Epistemology and Inquiry: the Primacy of Practice', in S. Hetherington (ed.), *Epistemology Futures,* Oxford, Oxford University Press.

Hume, D. (1975), *Enquiry Concerning Human Understanding*, in L.A.Selby-Bigge (ed.), *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, Oxford, Clarendon Press.

James, W. (1948), 'The Will to Believe', in *Essays in Pragmatism,* New York, Hafter Publishing Co.

MacIntyre, A. (1980), *Whose Justice? Which Rationality?*, Indiana, University of Notre Dame Press.

McCulloch, G. (1999), 'From Quine to the Epistemological Real Distinction', *European Journal of Philosophy* 1, (1): 30-46.

McGinn, C. (1977), 'Charity, Interpretation and Belief', *Journal of Philosophy* 74, (9):521-535.

Mele, A. (2001), *Self-Deception Unmasked*, Princeton, Princeton University Press.

Montaigne, M.de (1958), *Essays*, Harmondsworth, Penguin.

Morton, A. (2004), *On Evil*, London, Routledge.

Orwell, G. (1984), *Nineteen Eighty-Four*, Oxford, Clarendon Press.

Plantinga, A. (1993), *Warrant: the Current Debate*, Oxford, Oxford University Press.

Quine, W.V.O. (1953), 'Two Dogmas of Empiricism', in his *From a Logical Point of View*, Cambridge (Mass.), Harvard University Press.

Quine, W. V. O. (1960), *Word and Object*, Cambridge (Mass.), MIT Press.

Rescher, N. & Brandom, R. (1979), *The Logic of Inconsistency*, Oxford, Basil Blackwell.

Ross, L., Lepper, M. R. & Hubbard, M. (1975), 'Perseverance in Self-Perception and Social Perception: Biased Attributional Processes in the Debriefing Paradigm", *Journal of Personality and Social Psychology*, 32: 880-892.

Ross, L. & Anderson, C.A. (1982), 'Shortcomings in the Attribution Process: On the Origins and Maintenance of Erroneous Social Assessments', in D. Kahneman, P. Slovic & A.Tversky (eds.), *Judgment Under Uncertainty: Heuristics and Biases,* Cambridge, Cambridge University Press.

Shapin, S. (1994), *A Social History of Truth*, Chicago, Chicago University Press.

Simon, H. (1957), *Models of Man: Social and Rational* , New York, Riley.

Stich, S. (1981), 'Dennett on Intentional Systems', *Philosophical Topics*, 12 (1): 39-62.

Strawson, P.F. (1962), 'Freedom and Resentment', *Proceedings of the British Academy* 48: 187-211.

Tversky, A. & Kahneman, D.(1981), 'The Framing of Decisions and the Psychology of Choice'*, Science*, 211: 453-458

Velleman, J.D. (2000), *The Possibility of Practical Reason*, Oxford, Clarendon Press.

Werth, L.F. & Flaherty, J. (1986), 'A Phenomenological Approach to Human Deception', in R.W.Mitchell & N.S.Thompson (eds.), *Deception: Perspectives in Human and Non-Human Deceit*, Albany, State University of New York Press.

Wikforss, Å. (2001), 'Semantic Normativity', *Philosophical Studies*, 102: 203-226.

Williams, B. (1973), 'Deciding to Believe', in his *Problems of the Self* , Cambridge, Cambridge University Press.

Winch, P. (1958), *The Idea of a Social Science,* London, Routledge and Kegan Paul.