

Experience, belief, and the interpretive fold

Tim Bayne
University of Oxford and St. Catherine's College
Manor Road
Oxford OX1 3UJ
United Kingdom
tim.bayne@gmail.com

Elisabeth Pacherie
Institut Jean Nicod UMR 8129
1 bis, avenue de Lowendal
75007 Paris
France
pacherie@ens.fr

This paper was published in (2004), *Philosophy, Psychiatry, & Psychology*, 11/1: 81-86.
Please consult the published version for purposes of quotation.

We see two clusters of questions arising out of the papers in this issue. The first cluster concerns the role of experience in the explanation of delusions. Do abnormal experiences play a central role in accounting for delusions, or are they at best only marginal? And if experience plays a central role in accounting for delusions, exactly what role does it play? The second cluster of questions concerns the interpretability of delusions. Are delusional utterances meaningful, or are they mere noises? And if they are meaningful, what makes it the case that they are meaningful?

1. Empiricism and rationalism

In our target paper (this volume) we followed Campbell in distinguishing empiricist (bottom-up) accounts of delusions from rationalist (top-down) accounts. An account is empiricist if it grounds the delusion in an abnormal experience of some kind; it is rationalist if it refuses to ground the delusion in an abnormal experience. On our reading of the literature, most recent accounts of delusions – particularly monothematic delusions – are broadly empiricist (for reviews see Langdon and Coltheart 2000; Davies et al 2001). We took Campbell to be running against the current wave of empiricist enthusiasm in

defending a version of rationalism according to which delusions arise directly from organic malfunction: according to Campbell, the abnormal experiences of delusional patients ought to be explained in terms of their delusional beliefs rather than vice-versa. In response, we argued that there is much to be said in favour of empiricist enthusiasm, at least with respect to certain monothematic delusions. (There would seem to be little prospect of giving an empiricist analysis of the delusion of believing that one's left ear is a second fertile womb, for example.)

A number of the contributors to this issue are less enamored with the empiricist approach. Hohwy suggests that plausible models of delusions should be both top-down and bottom-up, while Broome claims that most contemporary models of delusion concur with Campbell's rejection of empiricism. Both Hohwy and Broome cite Frith's model of delusions of alien control as a non-empiricist model, although they present different accounts of how it is non-empiricist. Should we rethink our commitment to empiricism?

We think not. We suggest that our critics have misunderstood what we meant by the terms "top-down" and "bottom-up". In fact, we think that there is a lot of miscommunication in this debate. So the first order of business is to get clear on the terminology.

Begin by noting that a bottom-up approach is bottom-up from *experience*. Bottom-uppers need not – and in fact do not – claim that cognitive biases and background beliefs play no role in the genesis of delusions. Cognitive biases can exert their influence in one (or both) of two ways: either pre-perceptually (by influencing experiential content), or post-perceptually (by influencing the beliefs that the patient forms on the basis of their experiences). The first (pre-perceptual) role is clearly in accord with the empiricist approach, and the second (post-perceptual) role for cognitive biases is perfectly consistent with the empiricist view that the beliefs in question have their primary ground in abnormal experiences.

There are two ways in which empiricist (bottom-up) accounts of delusions can proceed (see Davies et al, 2001). According to what we called the *endorsement* version of empiricism, the person experiences that P and comes to believe that P. By contrast, *explanationist* versions of empiricism hold that the person simply has a strange experience, and comes to believe that P in an attempt to make sense of this experience. Both models are empiricist in that they hold the primary ground of the delusion to be an abnormal experience of some kind, but only proponents of the endorsement model identify the content of this abnormal experience with the content of the patient's delusion.

The distinction between endorsement and explanationist models is orthogonal to the distinction between one-factor and two-factor versions of empiricism (see Davies et al, 2001). One-factor versions of empiricism, such as Maher's (1999), hold that the only impairments delusional patients have are perceptual: their belief-fixation processes – that is, those processes that take perceptual states as input and generate doxastic states as output – operate within the normal range. Two-factor accounts, by contrast, hold that delusional patients have abnormal belief-fixation processes. It is a mistake to think that endorsement models just are, or must be developed in terms of, one-factor accounts (as Gerrans 2003, 50 seems to). An endorsement model can be developed in two-factor terms: normal individuals are able to inhibit the pre-potent response of believing what they perceive (think, for instance, of the Müller-Lyer illusion.) Perhaps some delusional patients have deficits in the inhibition of this pre-potent response. And explanationist models are compatible with both one-factor and two-factor accounts. Whether or not the explanationist needs to invoke a belief-fixation abnormality depends on whether she thinks that a normal individual would form (and maintain) the sorts of explanations of their unusual experiences that delusional patients do.

Does our distinction between empiricist and rationalist accounts map onto Klee's distinction between one-stage and two-stage accounts? Probably not. According to Klee, a one-stager holds "that "raw" perceptual experience contains its own intrinsic thematic content", while the two-stager holds that "thematic content is always supplied by a distinct stage of cognitive interpretation" (p.5). Klee might appear to be saying that a two-stage model holds that perceptual states lack representational content, but we suspect that that is not his view. (Although Klee does suggest that Davidson is a two-stager, and Davidson denies that perception has representational content.) Rather, Klee's two-stager seems to hold that perception inherits its representational content from the "belief-desire system". Although we are not entirely sure what it would be for a perceptual states to inherit its content from a belief-desire system, as far as we can tell this view is consistent with both empiricist and rationalist accounts of delusion-formation, as we (following Campbell) are using those terms.¹

¹ Klee goes on to say that some of Wittgenstein's holist views call into question the coherency of one-stage models of any kind of belief-state (9). How does that argument go? We're not sure. Klee says that beliefs are *born* networked together in complex relations of support and counter-support. That might be true – but we cannot see what that claim has to do with the question of whether or not perceptual states have their own ("raw") content.

So much for taxonomic clarification: let us return to the question of whether current models of delusion are typically empiricist or rationalist by focusing on Frith's model of alien control. Broome suggests that Frith's model is a top-down model, while Hohwy describes it as a mixed model. While there might be certain uses of "bottom-up" and top-down" on which these descriptions are correct, we think that Frith's model is pretty clearly a bottom-up model on our use of the term. Frith attempts to explain delusions of alien control in terms of the content of the patient's abnormal *experience* of agency.² And although he has said very little about how such patients move from these experiences to delusional belief, Frith clearly assumes that they form their delusional beliefs on the basis of these alien experiences (perhaps in conjunction with cognitive biases of various kinds).³ This approach to alien control is diametrically opposed that which a rationalist (such as Campbell) would take. A rationalist would say that if such patients have unusual experiences of agency, it is because they have delusions of alien control. Whether or not this rationalist account is plausible, it is certainly not supported by Frith's laboratory.⁴

To clarify this point, consider Hohwy's characterization of current models of the delusion of alien control. What Hohwy describes as a top-down effect – the modulation of the experience of self-initiated movement by hypofrontality – does not look at all top-down to us, for it plays a role in fixing *experiential* content. And what Hohwy describes as bottom-up – the fact that there is no inhibition of pre-potent doxastic response – does not look particularly bottom-up to us. Finally, Hohwy's discussion of predictive coding seems to be us to be grist for the empiricist mill. As far as we can tell, predictive coding is meant to play a role in the fixation of perceptual content, rather than account for the fixation of

² It is not clear whether Frith's approach is in terms of the explanationist model or the endorsement model. Klee presents him as an explanationist, but on occasions Frith appears to think that the experience underlying delusions of alien control has roughly the content of the delusional state itself – which would suggest an endorsement approach. Frith – like most empiricists – has not been particularly clear on what he thinks the contents of the experiences underlying the delusion of alien control are.

³ One may still suspect, as Klee does, that Frith's model does not explain why an anomalous experiences of agency due to defective predictive coding gives rise to delusions of alien control. But see Jeannerod and Pacherie (2004) for a way of supplementing Frith's account.

⁴ Although we lack the space to make this point here, we think that much the same goes for other models of delusion that Broome mentions. The models appeal to biases of various kinds, but they see these biases as operating on, or in conjunction with, abnormal experiences of some kind. In short, they are empiricist models in our sense of the word.

belief independent of perception. As such, models that appeal to it fall squarely within the empiricist camp.

We suspect that much of the apparent disagreement between Broome and Howhy on the one hand and ourselves on the other is verbal. In contrast, our disagreements with Klee are substantive. Klee rejects empiricist models on the grounds that they fail to explain why the delusions in question have the specific thematic contents that they do (4). His point seems to be not that the empiricists are unable to explain why delusional individuals have the abnormal experiences that they do, but that empiricists cannot explain why delusional patients form the delusions that they do given their unusual experiences. Why, Klee asks, does the unusual experiential state of Cotard sufferers lead them to form the belief that they are dead rather than a number, a large rock or a tub of dirt?

As Sass (this volume, 12-13) points out, deadness is a condition of a person and seems a natural choice for an individual who still has a subjective life, although stripped of its emotional component, and still experiences him- or herself as a person. There is no particular problem integrating Sass's explanation of why the Cotard delusion has the specific thematic content into an empiricist account. The Cotard patient retains background beliefs about his personhood together with current subjective experiences that cohere with these beliefs. According to an endorsement empiricist account, the Cotard's patient experiences himself as dead (or rotting, etc) as a result of his lack of affect and the pre-perceptual influence of his background beliefs. According to an explanationist account, the Cotard patient experience a strange and disturbing lack of affect, and adopts the belief that he is dead in order to make some sense of this experience. This seems to be a comprehensible explanation for the patient to adopt given the fact that he is aware that he is (or at least was) a person.

2. Interpretation

We turn now to questions of meaning and interpretation. Are delusional states interpretable? Can we bring delusional patients within the interpretive fold? There is much to be said here, but due to space constraints our comments will be programmatic.

There are three main tenets of Davidsonian interpretationism (Davidson 1975; Davidson 1994):

- (1) Rationality is constitutive of intelligibility: to explain behaviour is to locate it in a wider network of cognitive states and activities connected according to norms of procedural rationality.

(2) Rationality is constitutive of meaning: what gives a state its content is its place in a network of rationally connected propositional attitudes.

(3) Beliefs can be justified only by other beliefs.

For Davidson, claims (1) and (2) are but the two faces of the same coin. In contrast Klee sees them as dissociable. He argues against (2) but wants to keep (1). Against (2), Klee claims that since clinicians usually have little trouble identifying the content of even stark delusions, rationality cannot be constitutive of mental content. Yet, he maintains that despite having understandable content, stark delusions are nevertheless necessarily inexplicable.

On our reading, Klee's argument against (2) is partly based on (1) and (3). Beliefs are explained by their rational relations to other beliefs, but stark delusions cannot be located in such a network of relations – they do not cohere with other things that the deluded subject believes aright. Here, Klee seems to endorse Davidson's strictures on what kinds of relations (rational in what Gerrans would call the procedural sense) and *relata* (beliefs and only beliefs) can be appealed to in accounting for intelligibility.

We do not think that these strictures are justified. We reject (3) and think that (1) and (2) are acceptable only if suitably modified such that: (i) the *relata* that fix intelligibility and content also include non-propositional attitudes (such as experiences); (ii) intelligibility and content are determined not just by norms of procedural rationality but also by other principles of inference and cognitive biases; and (iii) intelligibility and content are not fixed in a globally holistic manner. Let us explore these three points briefly.

Despite its fame the Davidsonian claim that only beliefs can provide reasons is false. The perception as of a tree in front of one gives one a reason to believe that there is a tree in front of one. This reason is defeasible, of course, but it is a reason nonetheless (Pollock 1986). And whatever exactly the best account of mental state ascription turns out to be, it is exceedingly plausible that it will give an important role to experiential states (perceptual states, emotional states and so on).

Why has experience been neglected in accounts of interpretation? One reason is the widespread assumption, shared by Davidson, that experiential states lack representational content. Another reason, pointed out by Gerrans, is that Davidson equates norms of rational belief-fixation with norms of procedural rationality (the familiar principles of inference as found in deductive logic and probability theory). The equation is dubious when restricted to its home ground – that is, when conceived of as an account of the fixation of belief on the basis of prior belief – but it is even more problematic when applied to the fixation of belief on the basis of experience.

Klee's second reason for declaring stark delusions necessarily inexplicable is, he claims, Wittgensteinian. A stark delusion, Klee suggests, involves content that negates what stands fast for us. Here Klee seems to assume that we are all prisoners of a unique, shared language game or worldview, and that anyone who steps out of its bounds thereby becomes unintelligible. This assumption is unwarranted. As Wittgenstein himself says, 'the same proposition may get treated at one time as something to test by experience, at another as a rule of testing' (1969, §98). We participate in various language games, and what stands fast — framework propositions — varies from game to game. To take just one example, the rules that philosophers and scientists adhere to in their professional lives often differ from those they adhere to in their everyday lives.⁵

It may well be that if someone does not abide by the rules of a certain game — if he negates that which stands fast within it —, his utterances will not be explicable as moves in that game. But it does not follow that they are *necessarily* inexplicable, for they may well make sense in the context of another language game. And to spin out the metaphor, stark delusions may be telltale signs that a different language game is being played, one whose rules might be discernable. In the same way that the student of philosophy or quantum physics gradually uncovers the rules of the games they are learning, psychiatrists and clinical psychologists might be able to uncover the rules by which their delusional patients abide. Eugenie Georgaca's analysis of a dialogue between a clinician and a delusional patient provides a useful example of one way this can be done.

Moreover, as Sass and Broome point out, language games can overlap and interpenetrate in various ways (Sass, p. 10; Broome, p. 8). Although we may have failed to convey this in our target paper, we concur with Sass in thinking that an interpretation of a delusion need not 'preserve the conventional meanings of a word in a what would seem a rather exact or rigid way' (Sass, p. 6). Sass's discussion reminds us that the semantics of a word may reflect the fact that it is part of several intersecting language games (for death, the language games would include biology, the subjective realm of consciousness, religious beliefs concerning life after death, supernatural entities, etc.). In his use of words, a delusional patient may break the rules governing one of these intersecting language games while retaining the meanings these words have in other language games.

In our opinion, Davidsonian interpretationism overstates the uniformity and homogeneity of our network of cognitive states conceived of as beliefs held together by rational relations. And the use Klee makes of Wittgensteinian ideas seems to skew them in the

⁵ Similarly, as Sass notes (p.10), many practically minded and even scientifically inclined persons find it possible to believe in various religious doctrines.

direction of Davidsonian uniformity. Wittgenstein, however, insisted that the foundational status of beliefs is relative to a given language-game, and we participate in a variety of language games.⁶ The problem with homogeneous cognitive networks is that they are brittle: intelligibility is easily jeopardized when only one kind of relation holds states together. Understanding and intelligibility are more resilient in heterogeneous cognitive networks that involve a number of overlapping language games and a plurality of representational states. Taking into account the role played by experience – in particular, the role played by 'certain formal or structural aspects or pervasive infrastructures of experience' (Sass, 10) – allows us to make (partial) sense of delusions that the Davidsonian is forced to place outside the interpretive fold.

The empiricist approach attempts to bring the patient within the interpretative fold by broadening the range of possible personal-level explanations to include not just propositional states but experiences, such as moods, emotions, and perceptual states. The thought is that we can come to get a grip on what the patient believes – and why they believe what they believe – by having a greater appreciation of the contents of his experience. In this regard, the empiricist approach traffics in personal-level explanations: personal level phenomena — delusional beliefs — are explained in terms of other personal-level phenomena — experiential states. But – as Gerrans and Ghaemi remind us – empiricist approaches can also be supplemented by sub-personal explanations. Sub-personal accounts can be invoked to explain: (i) why the person had the experiences they did; and (ii) why they went on to form the beliefs that they did in response to those experiences. These accounts will be, at least in part, pitched at the systemic level, the biological level, or both.

References

- Campbell, J. 2001. Rationality, meaning and the analysis of delusion. *Philosophy, Psychiatry, and Psychology* 8: 89-100.
- Davidson, D. 1975. Thought and Talk. In S. Guttenplan (ed), *Mind and Language*. Oxford: Oxford University Press.

⁶ Broome (7-8) rightly stresses this point, but mistakenly assumes that it favours rationalist over empiricist accounts. On the contrary, an advantage of the empiricist account is that it promises to explain – in terms of systematic alteration of experience – the fact that a certain proposition has a framework status for a delusional subject.

- Davidson, D. 1994. Donald Davidson. In S. Guttenplan (ed), *A Companion to the Philosophy of Mind* (pp. 231-235) Oxford: Blackwell.
- Davies, M., Cotheart, M., et al. 2001. Monothematic delusions: Towards a two-factor account. *Philosophy, Psychiatry, and Psychology* 8: 133-58.
- Gerrans, P. 2003. A one-stage explanation of the Cotard delusion. *Philosophy, Psychiatry, & Psychology*, 9/1: 47-53.
- Jeannerod, M. & Pacherie, E. 2004. Agency, Simulation and Self-identification. *Mind & Language*, 19, 2: 113-146.
- Langdon, R. & Coltheart, M. 2000. The cognitive neuropsychology of delusions. In M. Coltheart & M. Davies (eds) *Pathologies of belief*. Oxford: Blackwell.
- Maher, B. 1999. Anomalous experience in everyday life: Its significance for psychopathology. *The Monist* 82: 547-70.
- Pacherie, E. & Proust, J. (to appear). Neurosciences et compréhension d'autrui. In E. Ennen, L. Faucher, P. Poirier & É. Racine (eds), *Des neurones à la philosophie: Neurophilosophie et philosophie des neurosciences*, Louvain-la-neuve: DeBoeck Université.
- Pollock, J. 1986. *Contemporary Theories of Knowledge*. Totowa, N.J.: Rowman & Littlefield.
- Wittgenstein, L. 1969. *On Certainty*, trans. D. Paul & G.E.M. Anscombe, eds. G.E.M. Anscombe & G. H. von Wright. Oxford: Blackwell.