

## Showing our Seams: A Reply to Eric Funkhouser.

*Abstract:* In a recent paper published in this journal, Eric Funkhouser argues that some of our beliefs have the primary function of signalling to others, rather than allowing us to navigate the world. Funkhouser's case is persuasive. However, his account of beliefs as signals is underinclusive, omitting both beliefs that are signals to the self and less than full-fledged beliefs as signals. The latter set of beliefs, moreover, have a better claim to being considered as constituting a psychological kind in their own right than the set of beliefs Funkhouser identifies.

When philosophers talk about beliefs, almost invariably they have in mind states that have the function of representing the world. When we think of beliefs along these lines, we naturally conceptualise them as deeply evidence-responsive states. We think of them as states with contents that are a function of the way the world is, in relevant respects, and which are designed to alter in ways that track changes in those worldly properties. Eric Funkhouser (2017) does us a valuable service in reminding us that some beliefs may have quite different etiologies. A range of evidence suggests that *by design* some of our beliefs have the contents they do for reasons that have less to do with representing the world and more to do with painting ourselves in a particular light to others. In this paper, however, I want to suggest that Funkhouser's account of beliefs as signals obscures almost as much as it reveals. First, it carves up the relevant territory in ways that are misleading, by dividing into different categories beliefs that trade off accuracy for other goods. Second, it misses two kinds of representational states that owe their existence to the signalling role: beliefs as signals to the self (rather than other agents) and what I will call fledgling beliefs. The latter deserve special attention, because unlike those he focuses on, these beliefs may have unique properties explained by their signalling role and may, therefore, deserves a place in our fundamental taxonomy of mind.

### *Funkhouser's Account.*

Funkhouser distinguishes two different kinds of functions that beliefs might have. One function is to represent the world as accurately as possible, in the service of guiding adaptive behavior. This is, he suggests, the primary function of beliefs. Beliefs are *for* navigating the world. In most species, this is probably their only function. But in some

species – perhaps only those with the capacity to attribute beliefs to others, and therefore perhaps only our own species – beliefs have another function as well or instead, of the navigational function. They are signals. To play this role, they need not accurately represent the world. Indeed, the aim of successful signalling can conflict with the aim of accurate representation.

A signal, for Funkhouser, is any object that has the function of communicating information to receivers, in order to modify the receiver's behavior. 'Function' is here understood expansively, encompassing not only biological design, but also cultural evolution and individual invention. Signals have the function of altering the beliefs and behavior of *other* agents. Of course, people sometimes *falsely* represent themselves as having beliefs in order to alter the behavior of others – that is, they lie. The kind of cases Funkhouser is interested in are cases in which agents signal beliefs they actually have: this signalling is honest signalling. Funkhouser argues that attending to certain features of beliefs, of agents and of cognitive mechanisms – their content, properties that indicate that they are designed to be easy to detect, properties that indicate honesty – suggest that some of our beliefs are best explained by this signalling function.

The explanatory payoff is explicitly etiological: Funkhouser aims to explain “the evolution or development of certain belief-forming biases or tendencies” (810). Beliefs that seem mysterious in the light of the navigation function, because they are predictably inaccurate, are explained by the signalling function. For instance, there is extensive evidence that we have certain self-serving biases. Consider the phenomenon of depressive realism. On some measures, at least, depressed people are more accurate in self-assessment than normal controls (Alloy and Abramson 1979; Alloy 1995). Having false beliefs about oneself is adaptive, and it may be adaptive because of the effects that the beliefs have on others. They may be better disposed toward us because we signal that we are better than we really are, for instance. Funkhouser also mentions pro-social beliefs and beliefs that contribute to group identity. He may have in mind here the evolutionary puzzle of altruism. This puzzle arises because it is mysterious how organisms could evolve a disposition to engage in behavior that benefits others (where 'benefit' is understood in terms of evolutionary fitness). Shouldn't such behaviors be selected against, and therefore go extinct, crowded out by more selfish strategies? There

are equally well-known responses to the puzzle, that cite reciprocity and kin selection, but they seem to fall short of explaining the extent of altruism in human beings. Funkhouser may believe – plausibly, to my mind – that the signalling function of beliefs helps to close the gap. By signalling a belief that we ought to act morally we signal our commitment to norms and thereby our willingness to cooperate with in-group members.

Why go to the trouble of actually believing the content we signal? Why not just lie, thereby getting the navigational benefits of accurate beliefs as well as the signalling benefits of manipulating others? In some cases, there is no mystery at all: the beliefs we signal are beliefs that accurately represent the world. But there are interesting cases – the cases Funkhouser aims to explain – in which there is a conflict between accuracy and signalling. When this occurs, we may be better off honestly signalling a false belief than lying, for a variety of reasons. One reason is that we happen to be set up in ways that make deception psychologically and practically difficult. We tend to be bad liars, and attempts at deception are often detected. Funkhouser is impressed, in particular, by Trivers' well-known evolutionary account of self-deception, according to which we deceive ourselves the better to deceive others (Trivers 2011; von Hippel & Trivers 2011). Further, there are psychological costs to successful deception: it takes a great deal of effort, and therefore entails cognitive load (which, in turn, reduces our capacities for other kinds of tasks). Moreover, there are high social costs to being caught in a lie, which act as a disincentive.

Funkhouser might boost his case by looking to the large literature in the cognitive science of religion on *credibility enhancing displays* (Henrich 2009) and *costly signalling* (Sosis 2003).<sup>1</sup> Credibility enhancing displays (in the context of religion) are behaviors that are reliably associated with genuine belief and therefore signal genuine commitment. They are often, but not always, costly signals (martyrdom for one's belief is a dramatic example; for more mundane cases, think of tithing, fasting, or even the opportunity costs associated with regular church attendance). The willingness of a person to engage in such costly activities is a signal of their sincerity. Typically, the costlier the display, the better a signal of commitment it is: an insincere believer would attempt to avoid paying such costs. But signals need to be hard to fake, rather than costly, to constitute credibility enhancing displays. While costly displays are hard to fake, not all hard to fake behaviors

are costly (Norenzayan 2013). Consider my willingness to cross a bridge over a high gorge. Crossing over may be in my interests, but my behavior enhances the credibility of my testimony that the bridge is safe.

### *Explaining Trade Off: Beyond Funkhouser*

Among the reasons Funkhouser gives for why it might be better to signal genuine but false beliefs than to lie is that such false beliefs may be otherwise beneficial. It is easy to see how this might be true. Having an inflated sense of one's own capacities (as we appear to do) might enable us to achieve goals we would not attempt were we more realistic. Believing that our children are smarter and intrinsically more deserving than those of our neighbors might motivate us to invest in them and to endure their more irritating moments. It is at this point that we run into the first problem for the theory: it is committed to distinguishing beliefs that are best classified together. Attention to how our beliefs benefit us directly – rather than in ways mediated by other agents – brings this out.

Recall that Funkhouser is interested in exposing and explaining a phenomenon that all too often escapes the attention of philosophers because they understand beliefs as states that have the functioning of guiding navigation. While that may be the *primary* function of beliefs, it is very hard to explain why we have some of the beliefs we do if we were concerned exclusively with getting things right. Instead, we seem to trade off accuracy for other goods; Funkhouser powerfully argues that one of these goods is the benefits that may accrue to us (or perhaps to our group) from signalling beliefs to others. But as Funkhouser also mentions, beliefs that are false can be useful for us for reasons other than the fact that they serve as signals to others. One – but only one – of those reasons may be because they serve as signals to *ourselves*, I suggest.

Funkhouser's signals are, explicitly, signals to other agents.<sup>2</sup> But there are cases in which sender and receiver are the same agent. We may signal to ourselves in order to convince ourselves more strongly (or at any rate to change our own dispositions). Positive affirmations are a staple of self-help books. There is extensive evidence, largely from sports psychology, that "self-talk" is effective (Tod, Hardy & Oliver 2011; Van Raalte,

Vincent & Brewer 2016). Interestingly, evidence for their efficacy when they are dishonest signals – when the person issuing them doesn’t believe them – is weak. But when they are honest signals, they appear to boost mood and may have longer term benefits (Wood, et al. 2009). These are honest signals of belief, made by ourselves to ourselves for the purpose of signalling.

Self-talk may work because we are social animals, and therefore give special weight to representations received via sensory transducers that are designed for the reception of external information. While we are selective in our acceptance of testimony, testimony received from ourselves is very unlikely to trigger filtering mechanisms since it does not usually evince the cues of untrustworthiness to which we are sensitive, such as membership of groups with which one does not identify, track record of malevolence to oneself or of epistemic unreliability (Mascaro & Sperber 2009; Sperber et al. 2010). We may also benefit from talking to ourselves for another reason: the architecture of the mind may impose certain limits on how information is shared via internal channels, limits that may be overcome by using different pathways. This suggestion is highly speculative, but there is some evidence in its favour, such as the finding that people better recall words they have spoken aloud than those they have merely read (Lafleur & Boucher 2014). As Dennett notes, a cognitive architecture with restricted pathways for internal sharing of information might encourage the habit of asking oneself questions aloud, or otherwise externalising thoughts, thereby blazing “a valuable new communicative trail between one’s internal components” (Dennett 2015: 45).<sup>3</sup> Self-talk may come to have the function of signalling our beliefs to ourselves, thereby changing our own behaviour. Perhaps some of our beliefs trade off accuracy for their success as signals to ourselves. There is no reason to exclude them from our taxonomy of beliefs as signals.

It would, of course, be easy for Funkhouser to modify his account to embrace self-signalling. There is a larger background issue, however, that is brought out when we focus on the ways in which accuracy and other goods may be traded off to benefit the self in ways that are not mediated by signalling to others. False beliefs may have benefits for us independent of *any* signalling function. Take our disposition toward the myside bias, which manifests in various ways (a disposition to recall evidence that supports hypotheses toward which are favourably inclined; a disposition to more easily detect

flaws in arguments for conclusions we are motivated to reject than arguments for conclusions we are motivated to accept, and so on). These dispositions may be adaptive (as I have argued elsewhere; AU 1), but they are adaptive because they conduce to the production of knowledge, not to reliable signalling. Signalling is a function that beliefs may play, but not one that must be invoked to explain the conflicts Funkhouser is interested in. We trade off accuracy for other goods for many reasons, not just signalling (to others or to ourselves) and a focus on signalling risks obscuring important commonalities between beliefs as signals and other beliefs.

Perhaps Funkhouser might object that his target is not belief-forming mechanisms but quite specific belief types. He aims to explain not the myside bias, say, but our predictable disposition toward thinking we're better than average or that our group is more noble than others, and so on. Once again, however, the explanation of such belief types is unlikely always to cite a signalling function. In fact, Funkhouser (unpublished) acknowledges that there are alternative explanations for many of the belief types he cites as evidence for his theory. For example, several theorists have argued that self-enhancing biases exist for the sake of mental health and the improvement of mood (e.g. Taylor & Brown 1988). While hedonic explanations probably fail (Mercier & Sperber 2017), other benefits to the self may well explain some of these biases (Sedikides & Alicke 2012). It is important to recognize that the extensive evidence that Funkhouser (unpublished) cites suggesting that many of these biases have a social function does not establish that that function is or involves signalling. If, as I have argued elsewhere (AU 1), biases can be adaptive for collective deliberation, we should expect them to be sensitive to cues of competition and social context even if they do not play a signalling function.

While Funkhouser maintains (plausibly) that at least some of the belief types mentioned probably have the (or *a*) function of signalling belief, he does not (and should not) suggest that the signalling account explains all of them. It does not even explain the existence of a class of beliefs that trade off accuracy for other goals. It picks out, from among our beliefs, some that have a particular etiology. It is true that this etiology has not sufficiently been attended to. It is also true that some of the beliefs that have this kind of etiology trade off accuracy for their signalling function, but not all the beliefs that trade off accuracy for other goods have a signalling function, let alone are signals to other agents.

For these reasons, it seems that the account has limited utility for cognitive science. The beliefs it picks out do not constitute a class in their own right.<sup>4</sup>

### *Fledgling Beliefs as Signals*

In the rest of this paper, I want to suggest that there may nevertheless be an interesting class of beliefs that owe their existence to their playing a signal function but which have a claim to constitute a class in their own right. This class consists of beliefs that Funkhouser excludes from his classification. Funkhouser is interested in what he calls “full-fledged beliefs” (827). What it takes for a belief to be full-fledged he (perhaps wisely) does not say. Rather, he limits himself to saying that “Whatever the nature of belief, I intend for beliefs-as-signals to meet those same conditions” (829; fn. 18). The really interesting and fundamental class of beliefs as signals, I want to suggest, are *not* full-fledged beliefs. It is because they are not full-fledged beliefs that they deserve our special attention. They have a better claim than the beliefs Funkhouser hones in on to constitute a fundamental kind.

What are the marks of paradigm or “full-fledged” beliefs? While the question is controversial, many philosophers would accept that such beliefs have the following properties:

- (1) Inferential promiscuity;
- (2) Evidence sensitivity;
- (3) Practical setting independence.

To say that beliefs are inferentially promiscuous (Stich 1978) is to say that they are apt to serve as premises in inferences with other attitudes with the right kind of content. My belief *that it is raining* may be used as a premise to infer the conclusion *that the picnic will be cancelled*. Evidence sensitivity is a matter of how beliefs are abandoned or modified in the face of evidence for or against them. I quickly abandon my belief *that the picnic will be fun* when I hear that Don plans to attend. Practical setting independence (Van Leeuwen 2014) refers to the property of governing behavior and inference across all contexts to which the content of the attitude is relevant. My beliefs about Don make me unlikely to invite him to a party, pretend not to notice him when he comes into view on the street and to avoid him at work and at the gym.

It is no mystery why beliefs tend to have these properties.<sup>5</sup> These are the properties that suit them to play their navigational role. A representation that responds to evidence (in the appropriate kinds of ways) is a representation that enables us better to track the changing show, and thereby negotiate our engagement with it more effectively. Inferential promiscuity also allows our representations to engage with each other in ways that allow us to integrate different sources of information, thereby raising the probability that we accurately track changes in the world. Practical setting independence ensures that whenever our behavior would be more attuned to the world by being guided by the representation, it will (tend to) be.

As Funkhouser argues, if the primary function of a belief is signalling to others, they may fall short of full possession of the kinds of properties characteristic of paradigm beliefs. In fact, their role as signals explains their “deviations” (815) from the dispositions or functional relations we expect of believers and beliefs. Such deviations may better enable them to play their signalling role. If it is in my interests to signal that I am healthy whether I am or not (because people will avoid me if I am unhealthy), then a degree of evidence insensitivity may be adaptive. If it is in my interests to signal that I am patriotic (because my neighbors may not trust me otherwise), and if I will more persuasively signal my patriotism if I am sincere (as Funkhouser suggests is generally true), it might be in my interests if the relevant representation does not engage inferentially with evidence that suggests that I am not all that patriotic after all (like my overseas investments, say). If it is in my interests to signal that I am well disposed to Don (because his sister is my friend, say), then perhaps I am better off if my representations of him are not practical setting independent, so that my distaste for him doesn’t show when I see her.

Funkhouser argues that the deviations from full possession of the properties characteristic of beliefs are too insignificant to cast doubt on their status as “full-fledged” beliefs. Their signalling function explains certain deviations from the properties we tend to associate with belief, but the deviations are relatively subtle. Some of these beliefs – that we are better than average, say – may be regarded as departing from the property of evidence sensitivity (even – *especially* – Don thinks he is better than average). Nevertheless, these departures are constrained, and for good reason: a belief that is too



divergent from obvious evidence won't serve as an effective signal to others since it is blatantly false. Moreover, these beliefs may have the other properties of the paradigm: they may be inferentially promiscuous and exhibit practical-setting independence.

But in addition to this set of full-fledged beliefs as signals, there may be a set of representations which – because they play a signalling role – deviate too far from possession of the properties characteristic of beliefs to count as full-fledged. I will call these representations *fledgling* beliefs (though there is no suggestion that they will ever become full-fledged). They deviate from full-fledged not only in evidence sensitivity and inferential promiscuity, but also in a property Funkhouser does not discuss: they exhibit a striking degree of practical setting *dependence* (governing behavior in some contexts but not others). Religious and ideological beliefs are the most obvious examples of such beliefs as displays.

There is both anecdotal and experimental evidence that religious and ideological beliefs sometimes fall dramatically short of being paradigm beliefs. Religious beliefs, for example, may tend to control behavior only when they are salient to individuals. The normative beliefs of Christians have a much greater effect on their behavior on Sundays than on other days of the week (Malhotra 2008). Christians appear to consume less pornography than non-Christians on Sundays alone (Edelman 2009). Similarly, reminders of religion dispose Muslims (Duhaime 2015) and Hindus (Xygalatas 2013) to behave in ways that accord with their religious norms. But the practical-setting dependence of religious beliefs extends beyond norms. For example, Harris and Giménez (2005) and Astuti and Harris (2008) found that Spanish children and Vezo children and adults (respectively) were more likely to attribute continuing mental life to a recently dead person, in accordance with the tenets of their religion, when given a religious prime than following a secular prime.

While practical-setting dependence may be the most dramatic way in which these beliefs diverge from the paradigm of full-fledged beliefs, they may exhibit lack of evidence-sensitivity and inferential promiscuity as well. Religious beliefs are notoriously inoculated against refutation (Talmont-Kaminski 2013). They also exhibit encapsulation: think of the person who attributes his surviving an earthquake to God's hand without

apparently noticing or being bothered by the worry that others, no less deserving or faithful, died.

In short, there is extensive evidence that religious beliefs often fail to possess the properties of inferential promiscuity, evidence sensitivity and practical setting. But does this fact distinguish them from non-religious beliefs? Mightn't these failures be perfectly general facts about the representations that guide our behavior?<sup>6</sup> While I am sceptical that these failures are the province of any particular domain, where 'domain' is picked out by the content of beliefs (AU 2), and I suspect that the differences between representations in each of these properties are differences in degree, it may be that these differences are sufficiently large and sufficiently clustered (such that when a representation exhibits any of these properties to a smaller extent than most beliefs do, it will likely fall short on all of them) to justify us distinguishing different classes of beliefs: full-fledged and fledging.

Of course, all beliefs are subject to salience or priming effects. But some beliefs are more easily primed than others. Indeed, some may be chronically accessible, such that it takes effort to suppress them. Those beliefs we describe as intuitive tend to possess these properties to a higher degree than those that are unintuitive. There is extensive evidence that representations differ from one another in these properties. More intuitive representations may be substituted for less intuitive, for example, when the context renders both salient. There here of *theological incorrectness*, for example, which describe the substitution of intuitive representations for the properties officially ascribed to god by believers (Barrett 1999). Similarly, even those with college-level education in scientific concepts may substitute more intuitive representations, even in contexts in which one would expect the scientific concepts to be salient or primed (Shtulman 2006; Shtulman & Calabi 2013).

There can be little doubt then, that representations differ significantly in properties like practical setting independence: some must be primed to control inference and behavior, whereas others must be suppressed if they are not to do so. It may nevertheless be doubted whether there is a large class of representations that have the set of properties that are supposed to be characteristic of full-fledged beliefs. While we shouldn't idealize

our beliefs, I think that the claim that many representations that agents take themselves to believe do tend to govern behavior and inference quite generally is extremely plausible. Think of entirely mundane beliefs like “doors are usually the best way of leaving a room”. People do not tend to try to walk through walls or climb out the window when a door is available, not even when they are tired, stressed or under cognitive load. The chronic accessibility of a construct is typically thought to arise out of its frequent activation, especially activation across different contexts: we should therefore expect such accessibility from a great many mundane beliefs. Others may (arguably) be chronically accessible for evolutionary reasons: such accessibility might be developmentally canalized

We might hope to explain the differences between those beliefs that are chronically accessible and those that fall short of possession of the properties that characterize full-fledged beliefs by reference to nothing more than how often and how broadly they are activated. The fact that religious (and ideological) beliefs seem so easily trumped by others, even in contexts in which their content is salient (as in the experiments cited above on afterlife judgments) suggests that such an explanation may be insufficient. Current accounts of the origins of religious beliefs do not seem to offer any explanations. On the standard model, religious belief is a byproduct of dispositions selected for other functions; especially agency detection (Powell & Clarke 2012). A minority of dissenters from this account argue that religious belief is selected for, perhaps as a result of group selection rather than genetic selection (Wilson 2002). Whatever the merits of these accounts as explanations of the origins and maintenance of religious belief, however, none of them seem to explain why religious beliefs often seem to fall short of being full-fledged beliefs. An emphasis on their role as signals takes up this explanatory slack.

Here I follow, and extend, Kim Sterelny, whose work on beliefs as *displays* predates Funkhouser’s. As Funkhouser recognizes, religious beliefs may play an important signalling role: they may indicate a commitment to certain norms or a concern with supernatural punishment, thereby signalling that the person whose states they are is a good bet for (otherwise risky) cooperative enterprises. Sterelny recognizes that beliefs with this kind of function will tend to be subject to pressures toward having the properties of navigational states. For instance, he emphasises how we might display our

beliefs to commit ourselves publicly and thereby bring it about that we embed the representation in our cognitive architecture (analogously, I might announce I have given up smoking in order to stake my reputation and therefore bolster my resolve). This kind of pressure tends to ensure that our displayed states tend not to fall *very* far short of being full-fledged beliefs (Sterelny 2015). But because their principal function is display, and not navigation, they may lack “inferential integration” and be relatively insensitive to new experience (Sterelny forthcoming).

Sterelny is not very explicit in how and why beliefs as displays may lack inferential integration and evidence sensitivity, but it is easy to see how the story might go. If the function of my belief is to signal commitment, then I need be concerned with the inferential entailments of my belief only to the extent to which failures of inference call into question the genuineness of my belief. That, of course, may happen (we question the genuineness of a commitment to equality, for instance, if the person fails to draw obvious conclusions from it) but not every failure of inferential promiscuity runs such a risk (this is especially true with regard to private failures). In fact, there may be positive advantages to a lack of inferential promiscuity, inasmuch as it allows the faithful to affirm their commitment without noticing unpalatable implications (as in the earthquake example above). Equally (though Sterelny does not notice this fact), the display function may predict and explain practical setting dependence: a display need only be active in those contexts in which there is a (relevant) audience to receive the signal, not all the contexts in which its content is relevant. A selective disposition to neglect religious beliefs may protect them against refutation and allow us to guide our behavior by the light of conflicting naturalistic commitments (Van Leeuwen 2014).

As already noted, Funkhouser adduces extensive evidence that the beliefs we signal tend to be honestly signalled. We are often transparent to others, sincerely believed claims are more convincingly transmitted or advertised, the punishments meted out to deceivers are powerful disincentives, and so on. But *honestly signalled* beliefs need not be *full-fledged* beliefs. A belief may be honestly held and yet fail to exhibit inferential promiscuity, evidence sensitivity or practical-setting independence. Honesty, in the relevant sense, is primarily a property of the person’s conscious relation to her beliefs. I am sincere in reporting my beliefs about Don insofar as when I examine them I feel no

(significant) doubts or reservations. But inferential promiscuity, evidence sensitivity and practical-setting independence are properties that beliefs have whenever their contents are relevant, not just when we entertain them. My honesty might be called into doubt if I do not respond to what I take to be evidence against my belief or to draw what I accept are the conclusions that follow from it, but my failure to appreciate good evidence or to see these conclusions does not always cast doubt on my honesty.

This is clearest with regard to the property of practical setting (in)dependence. If I sincerely believe that  $p$ , then whenever my belief that  $p$  is occurrent it will dispose me to act consistently with it. But such consistency may require top-down processing: I may need to *draw* the right inferences, for example. Thus my belief may be sincere but not full-fledged: it is sincere because whenever I ask myself whether I believe that  $p$ , I experience no (or at any rate, little) doubt as to its truth. But the belief is not full-fledged because playing its signalling role does not require it to play the full range of functional roles characteristic of such beliefs.

Whereas the set of beliefs as signals on which Funkhouser focuses – full-fledged beliefs – share all their other interesting properties with other beliefs that trade off accuracy for other goods, fledgling beliefs may have a stronger claim to belonging to our fundamental taxonomy. The combination of practical-setting dependence, lack of inferential integration and encapsulation from evidence they manifest may be unique. This combination is explained by their role as signals, rather than as navigational states.

Perhaps, however, these states are not beliefs at all, and whatever their intrinsic interest they do not belong in our taxonomy of such states. So Funkhouser (who is after all very well aware that there are signals that are not beliefs), might object. Such an objection has precedent and is not without plausibility. Schwitzgebel (2001; 2010) for instance argues that states that have associated with them some but not all of the dispositions constitutive of full-fledged beliefs should be regarded as in-between beliefs; perhaps (contra Schwitzgebel), such states should not be treated within an account of belief. There are, however, good reasons to think that fledgling beliefs as signals are nevertheless genuine beliefs. One reason is that they are professed by the agent: they are taken to be beliefs. Frequent assertion is the single most powerful cue for belief attribution by ordinary

people, overwhelming evidence of behavioural (in)consistency (Rose, Buckwalter & Turri 2014). It is not mere respect for first-person authority that accords our professed beliefs the (defeasible) status of genuine belief. It is the fact that our professed beliefs do tend to control behavior, and exhibit evidence sensitivity and inferential promiscuity, at least *when they are occurrent*. They have the hallmarks of belief much of the time. They may even tend to control our behavior across *most* settings, because there is a high correlation between cues that remind us of them, thereby causing them to be occurrent, and the relevance of their content.<sup>7</sup>

Perhaps we may discover that when beliefs as signals fail to have (pretty much) *all* the properties of full-fledged belief they tend to have *few* of the properties of such beliefs. Perhaps, that is, such states are too different from full-fledged beliefs to be beliefs at all. On this view, there are few or no fledgling beliefs; belief is a taxon, rather than a stretch of a continuous set of dimensions. If 'belief' picks out such a stretch, then we should expect there to be some vagueness as to what gets included, and there is room for fledgling beliefs as well as full-fledged. There is extensive evidence that suggests that representational states do vary continuously: while there are pressures for beliefs to come to have the role of navigational states, many of our apparent beliefs fail to occupy these roles in sometimes surprising ways (see Bortolotti 2009 for discussion). There seems plenty of room for fledgling beliefs; beliefs that are professed but play their navigational role too imperfectly to be full-fledged.

If belief is continuous, then perhaps there are no kinds to be delineated non-arbitrarily. The difference between fledgling and full-fledged beliefs is a difference in the degree to which they manifest properties that vary continuously and any way of dividing up the territory here would reflect a terminological decision. That is an open possibility right now. However, there are reasons to suspect that though the properties may vary continuously, we will find that they tend to cluster (perhaps homeostatically). States that are apt for navigation tend to have the properties of sensitivity to evidence, aptness for inference and practical setting independence because these properties conduce to one another. These states tend to cluster up one end of the continuum, varying from one another in the possession of these properties to be sure, but within a narrow range. We may find that states that are apt for display come under these pressures to a lesser extent

and that those states that depart from these properties – and therefore occupy a broader range of the territory – tend to be display states, not control states. If that's right (and of course this remains speculative), then we might want to identify these states as a distinct kind. They may nevertheless qualify as beliefs, because they are professed and because when they are occurrent they play the navigational role.

### *Conclusion*

Funkhouser has done us a service by reminding us of a fact that is too often forgotten, and which indeed may come as news to many: that beliefs may have aims other than truth. His article illuminatingly set out the conditions for qualifying as signals, and argues persuasively that some of our beliefs have this function. It also demonstrates that bearing this function in mind helps to explain some otherwise puzzling features of some of our beliefs. For all this, he deserves our gratitude.

But his account excludes other representations that qualify as beliefs – some fledgling and some full-fledged – and which play a signalling function. Some of these beliefs are signals not to others but to ourselves. Some of them fail to be paradigm beliefs *because* they are signals, departing from the profile of such beliefs in ways Funkhouser fails to consider. Funkhouser therefore misses the class of beliefs as signals – fledgling beliefs – that has the best claim to inclusion in our taxonomy of representations in its own right.

We may yet find that fledgling beliefs are too heterogeneous to constitute a kind. They may not even constitute a historical kind: perhaps beliefs that fall short of being full-fledged on the properties of integration, evidence sensitivity and practical setting-independence have a variety of origins (in which case my points here would constitute more in the way of friendly amendments to Funkhouser's view – calling attention to implications he overlooks and evidence he neglects – than a serious objection to it). Or we may find that belief is not sufficiently dichotomous for us to mark off kinds in the way I have suggested. Like Funkhouser's, my paper leaves us with more questions than we had before. Perhaps that's a sign of progress.<sup>8</sup>

### NOTES

<sup>1</sup> In the as-yet unpublished companion piece to “Beliefs as Signals” which he kindly shared with me, Funkhouser mentions costly signals but not the more encompassing notion of credibility enhancing displays.

<sup>2</sup> In his general account of signals (810-14), Funkhouser does not stipulate that the receiver is another agent. That account has the flexibility to accommodate self-signalling. But in his account of beliefs as signals (815), he is explicit that the function of the signal is to alter the beliefs and behaviors of *other* agents.

<sup>3</sup> Carruthers (2015: 164) suggests that the same effect – the recruitment of “all of the brain’s circuits to evaluate it, and to compete with one another in offering suggestions about what should come next” – can be achieved by inner speech as effectively as outer. There may be grounds for regarding such internal self-talk as a signal to the self, too.

<sup>4</sup> It is true that attending to the signal function may sensitize us to certain properties of agents and cognitive systems we might otherwise miss. For example, Funkhouser shows how bearing it in mind leads us to look for reliable detection systems on the part of receivers and otherwise puzzling signs of translucency on the part of senders. This is an important lesson. But the lesson does not concern *belief* per se; indeed, it does not concern cognitive systems per se. Being on the look out for signals sensitizes us to features we might otherwise miss, but they are features of phenotypes (from fur to feathers to behavior to emotions), and beliefs should occupy a small part of our attention in this territory.

<sup>5</sup> Obviously, we should not idealise beliefs. None of our beliefs have these properties to the fullest extent. Either due to properties of the representation itself, of other features of our cognitive architecture, or due to simple performance errors, we may fail to draw an obvious inference or to govern our behavior in accordance with our beliefs. However, we do appear to have representational states – paradigm beliefs – that possess these properties to a high degree and our inference and behavior does tend to be governed by them under conducive conditions (when we are not under cognitive load or fatigued, for example).

<sup>6</sup> I owe this objection to a reviewer for this journal.

<sup>7</sup> A reviewer for this journal reminds me that explicit attitudes are themselves surprisingly poor predictors of behavior. Nevertheless, across domains in which we lack incentives to dissimulate our attitudes and we face few motivational obstacles, they appear to be much better than implicit (Perugini 2005; Oswald et al. 2013). Implicit representations explain only a very small part of the variance in our behavior (AU 4). Elsewhere I have argued that the kinds of implicit representations probed by the implicit association are not beliefs, on the grounds that they lack the properties of evidence sensitivity and inferential integration (AU 3). But implicit representations are *not* professed by the agent, and therefore do not control behavior in a belief-like way even when activated. That is, they do not play the role of control states for the purposes of effortful control. It is worth adding that part of the explanation for the surprisingly low predictive validity of explicit measures may be that in many of the domains in which attitude-behavior correlations have been probed – attitudes toward social and political issues, and racial groups – explicit attitudes may have a signalling function.

<sup>8</sup> I am extremely grateful to two reviewers for this journal, whose comments forced me to become very much clearer both regarding Funkhouser’s views and my own.

## References



Alloy, L. B. & Abramson, L. Y. (1979) The judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, 108, 441–85.

Alloy, L. B. (1995) Depressive realism: Sadder but wiser? *The Harvard Mental Health Letter*, 11, 4–5.

Astuti, R. & Harris, P.L. (2008). Understanding morality and the life of the ancestors in rural Madagascar. *Cognitive Science*, 32, 713–740.

Barrett, J L. 1999. Theological Correctness: Cognitive Constraint and the Study of Religion. *Method & Theory in the Study of Religion* 11: 325–339.

Bortolotti L. (2009). *Delusions and other irrational beliefs*. Oxford University Press, Oxford.

Carruthers, P. (2015). *The Centred Mind*. Oxford: Oxford University Press.

Dennett, D.C. (2015) *Elbow Room: The Varieties of Free Will Worth Wanting* (New edition). The MIT Press.

Duhaime, E. (2015). Is the call to prayer a call to cooperate? A field experiment on the impact of religious salience on prosocial behavior. *Judgment and Decision Making*, 10, 593–596.

Edelman, B. (2009). Markets: Red Light States: Who Buys Online Adult Entertainment? *Journal of Economic Perspectives* 23, 209–220.

Funkhouser, E. (2017). Beliefs as signals: A new function for belief. *Philosophical Psychology*, 30(6), 809–831.

Funkhouser, E. (in press). Beliefs as signals: Examples and empirical support.

Harris, P., & Giménez, M. (2005). Children's acceptance of conflicting testimony: The case of death. *Journal of Cognition and Culture*, 5, 143–164.

Henrich, J. (2009). The evolution of costly displays, cooperation, and religion: Credibility enhancing displays and their implications for cultural evolution. *Evolution and Human Behavior*, 30, 244–260.

Lafleur, A. & Boucher, V.J. (2014). The ecology of self-monitoring effects on memory of verbal productions: Does speaking to someone make a difference? *Consciousness and Cognition*, 36, 139–146.

Malhotra, D. K. (2008). (When) are Religious People Nicier? Religious Salience and the 'Sunday Effect' on Pro-Social Behavior. Harvard Business School NOM Working Paper No. 09-066.

< <http://dx.doi.org/10.2139/ssrn.1297275> >

Mascaro, O. and Sperber, D. (2009). The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition*, 112, 367–80.

Mercier, H. & Sperber, D. (2017). *The Enigma of Reason*. Harvard: Harvard University Press.

Norenzayan, A. (2013). *Big Gods: How Religion Transformed Cooperation and Conflict*. Princeton: Princeton University Press.

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., and Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: a meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, 105, 171–192.

Perugini, M. (2005). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology*, 44, 29–45.

Powell, R., & Clarke, S. (2012). Religion as an evolutionary byproduct: A critique of the standard model. *The British Journal for the Philosophy of Science*, 63, 457–486.

Rose, D., Buckwalter, W. & Turri, J. (2014). When Words Speak Louder Than Actions: Delusion, Belief, and the Power of Assertion. *Australasian Journal of Philosophy*, 92, 683–700.

Schwitzgebel, E. (2001). In-Between Believing. *Philosophical Quarterly*, 51(202), 76–82.

Schwitzgebel, E. (2010). Acting Contrary to Our Professed Beliefs or the Gulf Between Occurrent Judgment and Dispositional Belief. *Pacific Philosophical Quarterly*, 91(4), 531–553.

Sedikides, C. & Alicke, M. (2012). Self-Enhancement and Self-Protection Motives. In *Oxford Handbook of Human Motivation*, ed. R.M. Ryan (New York, NY: Oxford University Press).

Shtulman, A. 2006. Qualitative differences between naïve and scientific theories of evolution. *Cognitive Psychology* 52: 170–194.

Shtulman, A., & Calabi, P. 2013. Tuition vs. intuition: Effects of instruction on naïve theories of evolution. *Merrill-Palmer Quarterly* 59: 141–167.

Sosis, R. (2003). Why aren't we all Hutterites? *Human Nature*, 14(2), 91–127.

Sperber, D. Clément, F, et al. (2010). Epistemic Vigilance. *Mind & Language*, 25, 359–393.

Sterelny, K. (2015). Content, Control and Display: The Natural Origins of Content. *Philosophia*, 43, 549–564.

Sterelny, K. (2017). Religion Re-Explained. *Religion, Brain & Behavior*, forthcoming.

Stich, S. (1978). Beliefs and subdoxastic states. *Philosophy of Science*, 45, 499–518.

Talmon-Kaminski, K. (2013). *Religion as Magical Ideology: How the Supernatural Reflects Rationality*. Durham: Acumen.

Taylor, S. & Brown, J. (1988). Illusion and Well-Being: A Social Psychological Perspective on Mental Health. *Psychological Bulletin*, 103(2), 193–210.

Tod, D., Hardy, J. & Oliver, E. (2011). Effects of Self-Talk: A Systematic Review. *Journal of Sport and Exercise Psychology*, 33, 666-687.

Trivers, R. (2011). *The folly of fools*. New York, NY: Basic Books.

Van Leeuwen, N. (2014). Religious Credence in not Factual Belief. *Cognition*, 133, 698–715.

Van Raalte, J. L., Vincent, A. & Brewer, B. W. (2016). Self-talk: review and sport-specific model. *Psychology of Sport and Exercise*, 22, 139-148.

von Hippel, W. & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences*, 34(01), 1–16.

Wilson, D.S. (2002). *Darwin's cathedral: evolution, religion, and the nature of society*. Chicago: University of Chicago Press.

Wood, J.W., Perunovic, W.J.E., & Lee, J.W. (2009). Positive self-statements: Power for some, peril for others. *Psychological Science*, 20(7), 860-866.

Xygalatas, D. (2013). Effects of religious setting on cooperative behaviour. A case study from Mauritius. *Religion, Brain and Behavior*, 3, 91-102.