

Book Review

Bayne, T., & Fernández, J. (Eds.). (2008). *Delusions and self-deception: Affective and motivational influences on belief formation*. Hove, UK: Psychology Press. Pp. 310. ISBN 978-1-84169-470-2. Price: £36.95 (hbk).

The multiauthor, multitopic book is the curse of the academic reviewer who often reverts to a restaurant-style review where each of the offerings is listed with a brief description and a comment on their subjective quality. Not being much of one for restaurants (or, indeed, lists) the recent book *Delusion and Self-Deception*, edited by philosophers Tim Bayne and Jordi Fernández, poses something of a challenge for me. Instead of a blow-by-blow account, I shall start with a general impression, drawn from self-declared prejudices, and will end by tackling some of the underlying theoretical currents, more likely drawn from my unrecognised biases.

First, the breadth of the volume is one of its strengths as it looks at classical psychiatric delusions as well as anosognosia and hysteria in an attempt to explore the relationship between recent philosophical accounts of self-deception and pathological beliefs. My only criticism in this regard would perhaps be one of the field in general, which has a rather nagging preoccupation with the Capgras delusion that far outweighs its prevalence and, perhaps more to the point, its representativeness of delusions in general. The discussions are clearly influenced by Alfred Mele's (2001) recent "deflationary" account of self-deception in which Mele disavows the folk psychological definition (which suggests we deliberately "trick ourselves" into believing something that we know isn't true) and instead reframes self-deception as involving situations where we come to believe something implausible because we are motivated to do so, largely through selective evidence gathering. Mele's account is an ideal fit for current theories of delusions because, in many ways, it is almost a nonpathological analogue of some psychological theories of delusion.

This is both a strength and a weakness and I must admit I have much sympathy for critics of the Mele account (e.g., Bermúdez, 2000) who suggest that it doesn't fully distinguish cases of self-deception from cases where one is simply wrong, even if it does provide a framework for integrating psychological constructs such as motivated reasoning and cognitive bias. Indeed, Mele's own chapter in the book tackles many of these processes

although those familiar with the psychological literature, from the likes of Kahneman, Tversky, Gilovich, and Ariely, may find it a little sparsely illustrated with experimental evidence.

Nevertheless, I can't help but think that Mele's account inadvertently highlights a problem not just with the conception of self-deception, but also with current theories of delusion; namely, that beliefs are often considered throughout the book to be binary entities where one believes (p) or does not ($\sim p$), something I call the Yoda problem after the Jedi master's clearly unrealistic admonishment to "do or do not, there is no try". Evidently, beliefs are held with differing levels of conviction, may be accorded differing levels of importance in reasoning and further belief formation, and may be variably accessible (indeed, as at least some beliefs supervene on memory, properties of memory become relevant for all general theories of belief). This is consistent with both evidence on normal belief and delusions but theories of delusions are still largely wedded to the Yoda model where belief is, against all current evidence, "delivered" to a global box of beliefs in an all-or-nothing fashion.

Against the extensive background of empirical evidence for the dimensional nature of beliefs, a model of self-deception could look quite different: Motivational factors could alter the importance or relevance of beliefs, so contradictory beliefs are either never activated (dare we say, consciously considered?) at the same time, or, if activated, the activation is so weak that it causes no significant cognitive dissonance and hence no motivation for the resolution of the conflict. This would also account for the case of "twisted self-deception", where people convince themselves of things they don't want to be true, such as that their partner is having an affair, because loss aversion can be a powerful motivator, as we know, for example, through extensive work on anxiety disorders.

Another tendency in the book is for authors to try to shoehorn phenomena into the "two-factor" model (e.g., Ellis, Young, Quayle, & de Pauw, 1997; Langdon & Coltheart, 2000) which, apart from also suffering the Yoda problem, is still clearly a preliminary approach and is based on the untested assumption that cognitive processes that contribute to the content of a delusion can be adequately separated from those that contribute to its maintenance. Nevertheless, as a tribute to the breadth of the book, there are contrasting views in Elisabeth Pacherie and Neil Levy's chapters, where they try and explicitly draw theory from the complexity of the phenomena rather than trying to fit the phenomena to the theory. Both approaches are, of course, important with the caveat that everyone needs to be working under the assumption that meeting in the middle to explain the same agreed-upon delusional phenomena is a useful goal.

Overall, the book clearly shows signs of this sort of constructive thinking both within the chapters and in its curating. The chapters on delusions as

traditionally conceived explore diverse and challenging possibilities and the chapters on anosognosia are an important reminder that pathologies of belief come in many types (and the reference to work showing that poor insight in anosognosia can be variable and situation dependent is a reminder that we need to think beyond global impairments). My interest was particularly piqued by the chapter on conversion hysteria potentially being a mirror-image of anosognosia, which I found both thought-provoking and original. After reading the book, I'm not sure that I have come away thinking that theories of self-deception are necessarily opening a new front in the understanding of delusions, but consideration of these topics in parallel has clearly engaged the authors with many rewarding results.

REFERENCES

- Bermúdez, J. (2000). Self-deception, intentions and contradictory beliefs. *Analysis*, 60(4), 309–319.
- Ellis, H. D., Young, A. W., Quayle, A. H., & de Pauw, K. W. (1997). Reduced autonomic responses to faces in Capgras delusion. *Proceedings of the Royal Society of London: Biological Sciences*, 264B(1384), 1085–1092.
- Langdon, R., & Coltheart, M. (2000). The cognitive neuropsychology of delusions. *Mind and Language*, 15(1), 183–216.
- Mele, A. (2001). *Self-deception unmasked*. Princeton, NJ: Princeton University Press.

VAUGHAN BELL

*Departamento de Psiquiatría, Universidad de Antioquia, Colombia
and Institute of Psychiatry, King's College London, UK*