

Introduction

Philosophy of Self-Deception

Patrizia Pedrini[†]
patpedrini@gmail.com

The phenomenon of self-deception is one of those topics that, perhaps more than others, is capable of intriguing and fascinating those who decide to devote to it a part of their studies and research. It is also a topic that, once encountered and reflected upon, does not leave us the same as before, in our relationships either with ourselves or with others. This can happen because we get in touch with the psychological event, which is pervasive and complex, and which we feel may have been crucial, for better or for worse, or at least insidious, at many junctures of our own existence. We sense that perhaps many decisions we made – maybe even more than we would be willing to acknowledge – have been made upon one variety or the other of self-deception – that is, upon beliefs that are false, that we additionally may, at times, have the sense that are false, and yet are strongly, sometimes even irresistibly wanted, or desired. Its disconcerting hallmark lies in the fact that we somehow seem to come to believe a proposition that we should at least doubt is likely to be true, and that we seem to do that because of a strong motivation to acquire that false belief. That is why self-deception is included among the so-called “motivated irrationality” phenomena, to which other phenomena also belong, e.g., wishful thinking, cases of precipitate believing under the influence of strong emotions, and so on.

It is thus easy to get caught up in the attempt to analyse it as to the best of our ability, so as to have a coherent description of it, and also a convincing explanation as to why human beings embark on it at all. It is also tempting to believe that, if we can come up with such a description, and such an explanation, we might perhaps be better equipped to identify its occurrence in

[†] Senior Research Fellow in Philosophy funded by the Mensa Society; Fixed-Term Professor, Dept of Philosophy, University of Rijeka; Assistant Fellow, College of Letters and Philosophy, University of Modena & Reggio-Emilia, and Dept. of Philosophy, University of Florence.

ourselves and others, and so, possibly, also try to overcome it. This may be the hope we might want to ascribe to those who believe that self-deception is not a good thing. Other people, however, consider self-deception bliss, by virtue of its allegedly evolutionary, or simply individual, advantages.

Although it was notably described by Donald Davidson¹, in the early days of the debate, as an *intentional* attempt at deceiving oneself, in the hope, among other things, of distinguishing it from other, non-intentional forms of motivated irrationality, many people subscribed later on to the *anti-intentional* view of self-deception promoted by Al Mele (2001), now also referred to as “motivationalism”, as Mele replaces the explanatory hypothesis of an intention to deceive oneself with a more palatable, paradox-free explanatory account in which a motivational state, mainly a desire, triggers self-deception and explains it convincingly. After Mele’s seminal work, the debate has flourished greatly, and many other related, and vital questions, the way to which was fully paved by Mele’s research and the subsequent discussion, have been tackled.

Many of these questions have been brilliantly addressed anew by the authors who have contributed to this issue, but other, brand-new ones have also been posed and argued for.

In his article “When Are We Self-Deceived?”, Al Mele provides a sketch of his view about how self-deception happens and, interestingly, he returns to the proposed set of jointly sufficient conditions for entering into self-deception and offers a couple of amendments.

Dion Scott-Kakures gets back critically to the traditional question of intentionalism; in his article: “Can You Succeed in Intentionally Deceiving Yourself?”, and argues that if we take the model of interpersonal intentional deception seriously, we ought to conclude that a self-deceiver, so regarded, deceives herself *unintentionally*.

Anna Elisabetta Galeotti (“Self-Deception: Intentional Plan or Mental Event?”) also addresses the issue of whether self-deception is an intentional plan or a mental event, and argues that self-deception is a complex mixture of things that we do and that happen to us; the outcome is, however, unintended by the subject, though it fulfils some of his practical, though short-term, goals.

¹ See Davidson 1985.

José Eduardo Porcher, in his “Against the Deflationary Account of Self-Deception”, critically examines the anti-intentional, deflationary strategy, where the theorist attributes to a subject just one belief – the false belief – as opposed to two beliefs, the true one and the false one, as supposed by intentionalists. He captivatingly suggests that the deflationary view contains a failure that support the neglected view that the self-deceived are not accurately describable as believing either of the relevant propositions.

Eric Funkhouser breaks into new territory, that of “Practical Self-Deception”, as his article is titled. He argues that, in the very same sense that we can be self-deceived about belief, we can be self-deceived about matters that concern our practical identities – e.g., our desires, emotions, values, and lifestyles –, and he offers an striking account of where practical self-deception is accommodated.

The thread of the practical issues concerning self-deception is also taken up by Carla Bagnoli, in her “Self-Deception and Agential Authority”, and by Dana Kay Nelkin in her “Responsibility and Self-Deception: A Framework”. Both of them go on to touch directly on specific moral questions raised by self-deception.

Bagnoli adopts a constitutivist approach to self-deception, which has the merit of explaining the selective nature of self-deception, as well as its being subject to moral sanction, while also describing it as a pragmatic strategy for maintaining the stability of the self, hence being continuous with other rational activities of self-constitution. However, she argues, its success is limited, and its costs are high: it protects the agent’s self by undermining the authority she has on her mental life.

Dana Kay Nelkin focuses instead directly on the question of whether and, if so, when people can be responsible for their self-deception and its consequences. In particular, she argues that a particular motivationist account, the “Desire-to-Believe” account, together with other resources, best explains how there can be culpable self-deception, and that self-deception is a good test case for deciding important questions about the nature of moral responsibility.

The “Desire-to-Believe” account is the target of my own contribution, “What Does the Self-Deceiver Want?”, where I argue that it is unlikely that the self-deceiver’s primary want to believe, or interest in believing that p , occurs as the result of a merely contingent interest in p being true, as one version of such

general account wants us to agree. I also assess various consequences of the view I favour, regarding the self-deceiver's avoidance behaviour, "twisted" self-deception, and whether we should provide a unifying explanation of "straight" and "twisted" self-deception, as we are encouraged to do by the Desire-To-Believe" account defenders.

Julie Kirsch, in her "Narrative and Self-Deception in *La Symphonie Pastorale*", addresses the ever-lasting sceptical issue of whether forging a personal narrative is always at risk of self-deception. She looks at the ways narratives can actually contribute to self-deception, but she argues that not all narratives are invariably self-deceptive. Rather, when they are not, they can make a very positive contribution to self-knowledge and moral understanding.

Mark Young ("The Therapeutic Value of Intellectual Virtue") argues that the development of intellectual character has necessary therapeutic value with regard to self-deception. A motivational/dispositional account of self-deception is offered and linked to a predominant psychological theory of virtuous character worked out by contemporary virtue ethicists and virtue epistemologists.

Lisa Bortolotti and Matteo Marnelli ("Self-Deception, Self-Delusion, and the Boundaries of Folk-Psychology") lead us directly into the domain of philosophical psychopathology as well as back to vital and more general philosophical issues, such as the psychological vocabulary we should use to capture and explain some specific mental phenomena, and argue that both self-deception and delusions can be understood in folk-psychological terms. They suggest that there is continuity between the epistemic irrationality manifested in self-deception and in delusion.

Massimo Marruffa ("Remnants of Psychoanalysis. Rethinking the Psychodynamic Approach to Self-Deception") gets back to how self-deception fits the crucial psychoanalytic topic of defence mechanisms. Building on Giovanni Jervis' criticism of psychoanalysis, he sets out to integrate that psychodynamic approach to defence mechanisms fully into the neurocognitive sciences.

In the "Commentaries" section, Clancy Martin and Alan Strudler focus on two texts: Kierkegaard's *Diary of the Seducer* and Shakespeare's *Much Ado About Nothing*, and use the phenomenon of seduction to explore the complicated philosophical and psychological terrain of how truth, trust,

deception and self-deception may interact in a process with which we are all intimately familiar.

Mark A. Wrathall offers an analysis of Sartrean “bad faith” and claims that it amounts to a motivated failure to apprehend the state of dis-integration that exists between one’s facticity and transcendence. This “failure to see” is explained by drawing on Merleau-Ponty’s account of perceptual ambiguity and existential opacity.

In the “Book Reviews” section, the reader will find Elisabetta Sirgiovanni reviewing Lisa Bortolotti’s *Delusions and Other Irrational Beliefs* (OUP, 2010), and Brad Bolman assessing Clancy Martin’s collection *The Philosophy of Deception* (OUP, 2009).

Last but not least, we have a “Interview” section, where Professor Amélie O. Rorty agreed to be interviewed by me and generously answered questions on how the self must be to be capable of self-deception, the adaptive fitness of self-deception, its motivational content, the failures of self-knowledge involved in self-deception, and confabulation, and on the lines of research on which she encourages self-deception theorists to embark.

The idea of compiling this issue dates back to July 2010, when I received the invitation to suggest a topic and a team of contributors. The help and encouragement I have had from the members of the editorial board from the outset has been incalculable; the enthusiasm I have encountered in all the contributors who agreed to write a paper and have subsequently been so generously ready to discuss their views with me and other referees unforgettable and immensely instructive. I thank each of the authors warmly for making this issue a busy “virtual lab” that has enabled me to reflect further on the topic. I also thank my diligent assistant, Alice Giuliani, for her decisive help in getting me into, and especially *out of*, the final editing.

REFERENCES

Davidson, D. (1985). Deception and Division. In E. LePore & B. McLaughlin (Eds.), *Actions and Events*. Oxford: Basil Blackwell, 138–148.

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