

- Self-deception involves some kind of motivated and biased belief manipulation. However, Mele denies that self-deception should be understood on the model of interpersonal deception.

- Mele distinguishes between the following two kinds of motivation for self-deception:

straight self-deception: Self-deception in which the agent deceives herself about something that she wants to be the case. (Typical)

twisted self-deception: Self-deception in which the agent deceives herself about something that she does not want to be the case. (Atypical)

- The phenomenon of self-deception can be captured: 1) lexically, 2) by examples, or 3) by theory (or some combination thereof).

- Beginning with a definition of ‘deceive’, Mele argues that we can be misled into thinking of self-deception on the interpersonal model. And Mele argues that this leads to 2 puzzles. Here are two assumptions about the meaning of ‘deceive’:

1. By definition, person A deceives person B (where B may or may not be the same person as A) into believing that p only if A knows, or at least believes truly, that $\sim p$ and causes B to believe that p.

2. By definition, deceiving is an intentional activity: nonintentional deceiving is conceptually impossible. (6)

These two assumptions generate the following 2 puzzles of self-deception:

static puzzle: How can a self-deceiver both believe that p and believe that not-p? (Or even worse, how can a self-deceiver believe that p and not believe that p?!)

dynamic puzzle: How can a self-deceiver intentionally deceive himself? How can the self-deceived simultaneously possess the sophistication to trick oneself, and the ineptitude to fall for these tricks?

Mele will avoid these puzzles by denying the assumptions.

- Let's say some more about motivationally biased belief manipulation. This manipulation can occur at various levels: evidence gathering, hypothesis formation, inference, or belief retention (compare this with the bottom of p. 11).

- Not all motivated irrationality (bias) is intentional, however. Mele presents the experimental study about how female caffeine drinkers react to a reported link between heavy caffeine intake and disease in women in order to illustrate this point. (12–13)

Some motivated irrationality is intentional, some is not. Contrast this point with the agency/antiagency dichotomy presented on p. 13. One might try to salvage the agency view, for example in accounting for self-deception, by appealing to unconscious intentions. Mele does not deny the possibility of such intentions, but he does not think that they are likely in self-deception. He recognizes that he needs to give an account of how the desire motivates without an intention.

- Note the 3 kinds of motivation that Mele distinguishes on the top of p. 18.

Q: Are the cases on p. 20 (e.g., Bob and the teenager) really ones in which there is no motive to believe?