

Self-Deception

What is Self-Deception?

Lying standardly involves (at least) two people: a subject *S* who lies, and a target *T* who is lied to. *S* intends, by asserting that some belief β is true, to make *T* believe that β is true, even though *S* knows β to be false. Thus, one person, who knows something, sets out to deceive another person, who does not.

Reasoning by analogy, self-deception would seem to involve the same person setting out to *deceive themselves* (and succeeding in doing so). But this seems impossible on the face of it. How can a single person play the roles of *T* and *S*, both sending and receiving a misleading message, both knowing that β is false and believing it is true? Much philosophical analysis has attempted to square this conceptual circle. Whether or not it succeeds is moot.

Despite the paradoxical nature of self-deception, everyday examples are easily envisaged. A persistent hard drug user boasts he can kick his habit whenever he wants. A student who keeps skipping class insists she wishes to pursue an academic career. A husband who habitually cheats on his wife swears that he loves only her. Here, actors stubbornly deny what observers can readily perceive: that their problematic deeds imply character flaws.

Such everyday examples highlight how self-deception is not merely accidental—a type of cognitive error. It is also motivated—a type of wishful thinking. In particular, the desire to maintain a positive self-view—to self-enhance—lies at the heart of much self-deception. To admit to a character flaw—to being, say, an addict, idler, or cad—entails a psychic cost, in the form of lower self-esteem or social shame. To avoid incurring that cost, the admission is avoided. Thus, every act of self-deception involves denial, a classic defence mechanism.

Self-deceptive denials range from moderate to extreme. Suppose one pays a sizeable sum to see a movie, but the movie disappoints. The conclusion that one wasted money, or was foolish enough to do so, is unwelcome. It can, however, be avoided by concluding that the movie was better than it was. Such rationalizations are convenient ways of resolving otherwise dispiriting cognitive dissonances. More unqualified forms of denial surface in clinical cases of narcissism. Here, the motive to self-enhance predominates so strongly that arrogant self-justification becomes routine and invincible.

What Mechanism Underlies Self-Deception?

Can the paradox of self-deception, outlined above, be resolved? If people do not literally lie to themselves—by dividing themselves into two distinct and antagonistic centres of consciousness—how does the process work psychologically?

One promising model draws an analogy with junk mail. Sifting through one's mailbox, one attempts to separate mail one wants to keep from mail one wants to discard. To do so effectively, it is not always necessary to open every envelope. Superficial inspection of the envelope itself—is it familiar and formal, or glossy and generic?—often suffices to reliably identify each item.

A similar dynamic may characterize self-deception. That is, self-deceivers may indeed suspect that their problematic deeds carry unflattering implications. However, they do not need to unpack that suspicion fully to be disturbed by it. Even “unopened”, it contains enough cognitive cues to signal that it is unwelcome. Accordingly, self-deceivers may opt not to dwell on its “contents”. The suspicion accordingly passes out of their minds, to be replaced by more reassuring reflections.

Formally put, the mechanism underlying self-deception may be this: more elaborate processing of information is short-circuited if less elaborate processing hints that it poses a potential threat, especially to one's positive view of self. Hence, there is no need to postulate two selves inhabiting one mind—one the deceiver, and the other the deceived. Indeed, there may be no need even to postulate the co-existence of two conflicting beliefs. An uncongenial belief may be nipped in the bud before it gets the chance to compete properly with a congenial one.

This junk-mail model is consistent several findings from experimental research. For example, people selectively forget critical (but not flattering) feedback that pertains to personality characteristics that they regard as of primary (but not of secondary) importance to their identity. This suggests that, on receiving the feedback that they tend to forget, people do not dwell on it as much, or do not process it as elaborately, as the less disturbing feedback that they tend to recall.

Can Self-Deception be Experimentally Demonstrated?

But let us step back for a moment. Examples of self-deception, of the sort provided above, are easily enumerated. However, they still involve some degree of subjective interpretation. Accordingly, they may be mistaken—like the farfetched diagnoses of Freudian psychoanalysis. The direct experimental demonstration of self-deception would therefore be welcome. Three classic examples are offered below.

In one study, participants initially had their self-esteem lowered or raised by informing them that they had performed either poorly or well on a bogus examination. Next, they listened to audio recordings of varying durations. Some featured their own voice, some other people's voices. Participants had to distinguish the one from the other. In addition, participants had their levels of physiological arousal concurrently assessed.

Now, hearing one's own voice normally increases arousal due to one's implicitly recognizing it. This was also the case here: increased arousal reliably indicated the playing of one's own voice. However, in some cases, participants mistakenly denied hearing their own voice when it was played, and in others, mistakenly affirmed hearing their own voice when it was not. Crucially, false denials were more common among participants who had had their self-esteem lowered, whereas false affirmations were more common among those who had had their self-esteem raised. This suggests that feelings of shame or pride made participants either reluctant or keen to identify themselves at a conscious level. Nonetheless, they still always identified themselves at an unconscious level. Here, participants seem to deny that they were experiencing particular perceptions.

In another study, participants were led to believe that two types of heart existed: the Type I sort, predicting longevity, and the Type II sort, predicting the opposite. They were also led to believe that the length of time they could keep their forearm submerged in a bath of ice water—either for a longer period or a shorter one—reliably indicated which type of heart they had. Upon taking the relevant test, participants acted so as to modify the amount of time they kept their forearm submerged in the direction that implied a greater likelihood of possessing the desirable Type I heart.

However, before taking the test, participants would either have possessed a Type I or a Type II heart. This fact would not have changed in the course of the experiment. Hence, it would have been pointless for participants to strive to submerge their arm for a longer or shorter period. That is, their performance on the test could only have been diagnostic of the type of heart they had, but could not have had any causal impact on it. Yet if participants "accidentally" found themselves keeping the hand submerged for a longer or short time, then

they would have been free to regard the reassuring test result as real. This is exactly what happened. Here, participants seemed to deny that they were acting on particular intentions.

A final study featured male participants who all claimed to be entirely heterosexual. They were divided into two groups: those who also reported being homophobic, and those who did not. Both groups watched three types of erotic movie clips: heterosexual, lesbian, and gay. As they watched, their penile circumference was continuously monitored with a plethysmograph.

Both homophobia and non-homophobic participants showed substantial and equivalent arousal to heterosexual and lesbian clips. However, the homophobic participants also showed mild arousal to the gay clips, whereas the non-phobic participants showed next to none. This pattern of findings is consistent with homophobia at least partly deriving from a latent bisexuality deemed morally unacceptable. Here, participants seemed to deny that they entertained particular desires.

Why Does Self-Deception Occur?

Self-deception entails irrationality. People cling to positive illusions about themselves, in the face of contradictory evidence, and despite the practical risks of misjudging themselves. Why would they do so? Several hypotheses have been advanced. Some contend that self-deception is a structural by-product of how the mind generally works. Others contend that self-deception is a functional adaptation facilitating survival and reproduction. Both classes of hypothesis are complementary rather than contradictory.

Structurally, self-deception may arise naturally out of the modularity of the human mind. If somewhat autonomous mental units each perform their own specialized tasks, then one would not expect them always to operate in concert. This view can be readily harmonized with the junk-mail model above.

Functionally, self-deception may assist the deception of others. Suppose—due to sex differences in reproductive biology—men generally seek to mate widely whereas women generally seek to mate wisely. If so, then men may evolve to deceive women about their fidelity, and women evolve to deceive men about their availability. As a counter-strategy, both sexes may then evolve to see through one other's deception. What next? One counter-counter-strategy would be for both sexes to deceive themselves into believing—at least at first—they are offering what the other sex really wants. By doing so, there would be no more deception to detect. Analogously, it may also be advantageous for people to deceive themselves about their intentions in other relationships, whether cooperative or competitive.

Aiden P. Gregg & Nikhila Mahadevan
University of Southampton

See Also: Bragging and Grandiosity; Cognitive Dissonance; Courtship, Deception in; Deception, Defining; Darwin, Charles; Denial; Disbelief, Suspension of; Evolution; False Signals; Freud, Sigmund; Narcissism; Over-Confidence; Psychoanalysis; Rationality; Self-Esteem; Self-Justification

Further Readings

Adams, H. E. Wright, L. W., & Lohr, B. A. Is homophobia associated with homosexual arousal? *Journal of Abnormal Psychology*, v. 105, (1996).

Gregg, A. P., Sedikides, C., Gebauer, J. E. "Dynamics of identity: Between self-enhancement and self-assessment." In S. J. Schwartz, K. Luyckx, & V. L. Vignoles (Eds.), *Handbook of identity: Theory and research*. New York, NY: Springer, 2011.

Greenwald, A. G. "Self-knowledge and self-deception: Further consideration." In M. S. Myslobodsky (Ed.), *The mythomaniacs: An inquiry into the nature of deception and self-deception*. Mahwah, NJ: Erlbaum, 1997.

Gur, R. C., & Sackeim, H. A. "Self-deception: A concept in search of a phenomenon". *Journal of Personality and Social Psychology*, v. 37, (1979).

Kurzban, R. *Why everyone (else) is a hypocrite: Evolution and the modular mind*. Princeton, NJ: Princeton University Press, 2010.

Mele, A. R. *Self-deception unmasked*. Princeton: Princeton University Press, 2001.

Quattrone, G. A., & Tversky, A. Causal versus diagnostic contingencies: On self-deception and on the voter's illusion. *Journal of Personality and Social Psychology*, v. 46, (1984).

Trivers, R. *The folly of fools: The logic of deceit and self-deception in human life*. New York: Basic Books, 2011.

(1,687 words: 1,571 text and subheadings; 116 cross-references and references)