

compos mentis

Self-Deception

August Smith

Wheaton College

ABSTRACT

In this essay I examine the possibility of bona-fide self-deception considered as an intellectual vice. I first briefly survey traditional construals of self-deception, concluding with Donald Davidson that self-deception necessarily consists of a set of contradictory beliefs that exist simultaneously in the mind through the presence of a sort of incorporeal "cognitive barrier." Like every vice, self-deception stems from epistemic pride due to a vicious distortion of one's pursuit of truth through deficient desires arising from a malformed will. This habitual failure of willpower often causes one to reject evidence that rebuts a view that they previously held or wished to be true—such repetitive, vicious ignorance produces this "cognitive barrier" that gives rise to self-deception. One's will might arrive at this deficient state through an inordinate desire for comfort or attempt to avoid the acknowledgement of an unpleasant truth, such as a cheating spouse or discreet alcoholism. Self-deception is opposed not only to the virtue of self-knowledge (as is *prima facie* true), but also strength of the will insofar as one's will is directly involved in a conscious avoidance of a self-deceptive cognitive state. Furthermore, while the vice of self-deception is inexorably related to vices of wishful thinking and willful naïveté, it is important to note that it is qualitatively distinct. Finally, self-deception does seem to come in degrees, and can become permanent through repeated failure of self-examination and exacerbation of a weak will. Ultimately, though all people are prone to self-deception, it can be avoided through a cultivation of a strong will and self-examination with the aid of others.

KEYWORDS

Epistemology, Self-deception, Virtue Theory, Vice, Cognitive Barrier

In this paper I will discuss the intellectual vice of self-deception, as well as its corresponding causal ancestry, characteristics, and methods of avoidance. When it comes to intellectual vices, it doesn't get much trickier than self-deception. For one, epistemologists do not agree on the necessary and sufficient conditions for legitimate cases of self-deception, nor do all even admit that there are instances in which bona fide self-deception is possible. For the purposes of this essay, I will assume that it not only possible (for reasons which I will explicate below) but rife in both past generations and contemporary society. John, in his first epistle, agrees. The Biblical author chides his audience: "If we claim to be without sin, we deceive ourselves and the truth is not in us" (1 Jn. 1:8). He argues that all people are influenced by the corruption of sin—thus, to believe that we are unaffected by such is to deceive ourselves, and to therefore be without the truth. Ultimately, like all intellectual vices, self-deception is a deficiency of intellectual virtue, consequently depriving one of some epistemic good; therefore, one should strive to avoid self-deception at all costs to promote a virtuous and flourishing intellectual life.

IS SELF-DECEPTION POSSIBLE?

The first task is to define self-deception. This duty, however, proves more difficult than initially expected. Self-deception is often popularly construed as simply "lying to oneself," yet this definition presents a problem: "to be self-deceived one must at some time have known the truth, or, to be more accurate, have believed something contrary to the belief engendered by the deception" (Davidson 2010, 4). To lie involves both believing a true proposition and expressing a proposition that is contrary to this true one with the intention to garner another's belief in such falsity. However, in the case of "lying to oneself," there is a clear inconsistency; how can one both know the truth of a proposition and simultaneously be self-engendered to believe a contrary proposition?

According to a more traditional unified and integrated sense of the self, this inconsistency proves fatal to the possibility of self-deception, as Amélie Rorty notes: "If the self is essentially unified or at least strongly integrated, capable of critical, truth-oriented reflection, with its various functions in principle accessible to, and corrigible by, one another, it *cannot* deceive itself" (McLaughlin 1998, 13, my emphasis). This internal, rationalistic coherentist picture of the self therefore disallows the prospect of legitimate self-deception. On the other hand, some conceive the noetic structure as a simple amalgamation of atomic

beliefs that “can be individually added, changed, and deleted without regard to their propositional environment” (Davidson 2010, 5). This model allows and even encourages intellectual inconsistency. However, this conception of the mind seems somewhat inept; the self is ostensibly more than a simple conglomeration of atomic beliefs.¹ Thus, a more appropriate way to conceive of the notion of self-deception is needed.

Others have postulated that self-deception be thought of in solely self-actualizing moral terms, as any sort of analysis of beliefs is entirely unhelpful. “[...] Self-deception cannot be eliminated by checking whether there are contradictions in my beliefs, for even if I were successful in that endeavor I may yet deceive myself, for the interest with which I approach these beliefs and this endeavor might be self-deceptive” (Strandberg 2015, 49). While I admit that we may deceive ourselves in intellectual analysis, if we carefully examine held beliefs, we may nonetheless come to realize present inconsistency. The simple fact that self-deception itself may often hinder our attempts to thwart the vice does not prevent the essence of self-deception from essentially consisting in an incompatible set of beliefs. Therefore, for the purposes of this paper, I will refer to self-deception as a case in which two or more inconsistent beliefs are held simultaneously as result of physiological, cognitive, or as in our case, habitual deficiency. In other words, self-deception consists of a “contradiction of beliefs.”

DEFINING SELF-DECEPTION

It remains to be answered, however, how two or more conflicting beliefs may be held simultaneously. Donald Davidson admits that this problem is significant but not insurmountable. He begins by asserting that it is indeed impossible for an intellectual agent to believe (p and not-p). For example, it is entirely incoherent for someone to say that they both believe in God and do not believe in God. Yet, on Davidson’s account, it is *not* impossible for an intellectual agent to believe (p) and (not-p) simultaneously (2010, 198). To explain this apparent contradiction, Davidson states “that people can and do sometimes keep closely related but opposed beliefs apart. To this extent we must accept the idea that there can be boundaries between parts of the mind” (211). He quickly reminds his readers that these are not literal psychological barriers, but rather conceptual objects to aid understanding.

1. See Dupuy 1998 for an extensive discussion of this issue.

compos mentis

How does this barrier affect the right functioning of the mind and promote self-deception? On this point it is worth quoting Davidson at length:

We should not think of the boundaries as defining permanent and separate territories. Contradictory beliefs [about a correlated subject] must each belong to a vast and identical network of beliefs about [this subject] and related matters if they are to be contradictory. Although they must belong to strongly overlapping territories, the contradictory beliefs do not belong to the same territory; to erase the line between them would destroy one of the beliefs. I see no obvious reason to suppose one of the territories must be closed to consciousness, whatever exactly that means, but it is clear that the agent cannot survey the whole without erasing the boundaries. (Davidson 2010, 211)

This model provides a helpful solution to the problem; an intellectual agent can hold belief (p) and simultaneously hold (not-p) if the two beliefs are part of an intricately related web of beliefs, but do not occupy identical cognitive "areas." This allows for a sort of intellectual barrier to arise (for reasons discussed below) which inhibits conscious cognitive contact between the two contrary beliefs, an event that would mandate the destruction of one. An agent whose mind has developed or does develop these cognitive barriers is one who is viciously self-deceived.

CAUSAL ANCESTRY

We have now completed the arduous task of defining the vice of self-deception. How then does this vice come about? Like all vicious habits, self-deception arises out of pride. When one begins to value subjective desires over the objective intellectual goods of truth and self-knowledge, intellectual self-deception may result. "The worry, rather, is that what people will or desire may produce patterns of belief formation that undermine their ability to weigh evidence, assess claims, and evaluate behavior (especially their own)" (Floyd 2004, 60). It becomes clear that the will plays an active role in the development of self-deception. Upon evaluating information and attributing probative weights to certain propositions, a pure, unadulterated desire for truth is almost never attainable. Our wills, rather, are often influenced by external and internal factors (arising out of pride) that

increase our personal intellectual prejudices and inhibit our search for truth through the development of these cognitive barriers.

I acknowledge that there are instances in which withholding the truth from *others* provides a greater benefit. Take, for instance, the German woman who lies to Nazi forces to prevent the horrific murder of the Jews hiding in her basement—certainly this is morally commendable. However, it is not unreasonable to claim that we should nonetheless pursue truth for *ourselves* in all circumstances. However, “truth and rationality aren’t the only things we value, of course, and often the other things we value influence our inquiry in such a way to make believing what’s true less likely” (Elshof 2009, 27). These values often cause us to fall prey to faulty inference, incomplete survey of available evidence, lack of thorough diligence, or sympathy in our perceptual evaluations (Davidson 2013, 42). Ultimately, self-deception results when we allow our wills to become subject to our own desires to the extent to which our cognitive abilities to effectively determine truth and apprehend contradictory beliefs are severely hindered, especially in regard to self-knowledge and assessing our own intellectual states.

CHARACTERISTIC THOUGHTS AND ACTIONS

These values may come about through various different processes. One of the most prominent ways in which a non-truth-conducive cognitive habit is formed occurs when one rightfully adopts a true belief based on their circumstances but refuses to relinquish or adapt said belief once circumstances change insofar as to not allow for the original belief to remain tenable (LePore & McLaughlin 1992, 30-31). Say, for example, Sheila rightfully believes that her husband is not cheating on her, as he has remained faithful to her throughout their marriage thus far. However, after a couple of years, empirical evidence begins to suggest otherwise. Her husband is constantly returning home later than usual, he is overly protective of his phone and email, and proves defensive and shifty in conversation. The evidence mounts—it is apparent that her husband is cheating on her, and most unbiased witnesses would agree if presented this data (Mele & Rawling 2010, 247). And yet Sheila maintains her now *false* belief that her husband remains faithful because she has habituated herself to believe what was once true—this causes a cognitive barrier to materialize in her mind. The belief that her husband remains faithful and knowledge that he is cheating are intricately related but

clearly inconsistent. Yet these beliefs do not “cognitively encounter” one another because of this barrier, and Sheila therefore remains in a state of self-deception.

What causes her to maintain this false belief amidst rising evidence to the contrary? In this instance it proves to be fear of coming to terms with a belief that will most likely cause significant pain and discomfort (Davidson 2010, 209). This fear demonstrates the vicious habituation of external values affecting the will—her concern for her own comfort (which stems from pride) causes her to miscalculate the evidence with which she is presented and consequently deceive herself. The same fear provides a motivation for vicious habituation in other characteristic actions of the self-deceived, of which addiction is one of the most common. One can easily imagine the alcoholic who believes that he only drinks in moderation—he implicitly fears that, if he rejects this false belief, he will be forced to acknowledge his addiction and (depending on his moral values) attempt to assuage his vicious habit, which will undoubtedly result in much distress.² This means he does not recognize his addiction, only serving to propagate his self-deception further (Coleman 2007, 4). His own pride utilizes self-deception to prevent his will from motivating him into a proper evaluation and confrontation of the inconsistent beliefs.

OPPOSING VIRTUES AND SUPPORTING VICES

Therefore, it is clear that self-deception is not solely opposed to self-knowledge, as is *prima facie* the case, but to strength of the will as well. Davidson explains: “An agent’s will is weak if he acts, and acts intentionally, counter to his own best judgement; in such cases we sometimes say he lacks the willpower to do what he knows, or at any rate believes, would, everything considered, be better” (Davidson 2013, 21). Furthermore, Roberts and Wood describe a properly functioning will (or willpower, rather) as that which moves one to act based on a proper construal of themselves and the situation at hand in accordance with their intellectual good (2012, 63). Self-deception inhibits this process by *improperly* construing the situation at hand, causing the will to fail to move one to appropriate action at a suitable time. Thus, self-deception is opposed to the virtue of a strong will and promotes the development of the vicious weak will.

One may claim that both cases above are simply illustrations of wishful thinking, rather than self-deception. This objection raises the question: which

2. For a thorough discussion of this claim, see Leeuwen 2009.

vices support and closely relate to self-deception? It is only fitting to start with wishful thinking. Wishful thinking and self-deception are not one in the same, as often construed. Rather, wishful thinking proves to be a species of self-deception, as it involves a mental state of self-deception that is achieved through specific circumstances: namely, when one is already in possession of knowledge that a said belief is false, but purposefully suppresses that knowledge in order to make room for a contrary belief that appears more attractive to him or her. For example, say Tom has failed his last seven math quizzes. He knows this fact and its propensity to discourage belief in a future passing grade, but that does not keep him from choosing to believe that he will, in fact, pass the next quiz, despite a complete failure to sufficiently prepare. Thus, it can be seen that wishful thinking is not a separate vice, but one that falls into the category of self-deception.

However, a vice that is closely related to but distinct from self-deception is willful naïveté. Thomas Aquinas, in his *Summa Theologiae* speaks on a related vice, which he calls “blindness of mind.” The Angelic Doctor states:

Sometimes it is due to the fact that a man’s will is deliberately turned away from the consideration of [the] principle [of intelligibility] according to Psalm 35:4, “He would not understand, that he might do well”: whereas sometimes it is due to the mind being more busy about things which it loves more, so as to be hindered thereby from considering this principle, according to Psalm 57:9, “Fire,” i.e. of concupiscence, “hath fallen on them and they shall not see the sun.” On either of these ways blindness of mind is a sin. (Aquinas 1969, II.2.15.1)

Aquinas, in this excerpt, explains that people often turn away from self-realization and intelligibility because of willful neglect, as well as absent-minded business. While this vice is starkly similar to self-deception, it is actually quite different; rather than consisting in a state of mind that concurrently holds to two contrary beliefs, willful naïveté involves a purposeful neglect of the dutiful pursuit of evidence which may in turn support a belief which one does not wish to hold. Finally, I must note that self-induced deception is not an instance of self-deception either. Self-induced deception, as distinct from self-deception, consists in consciously convincing oneself of a false belief, which does not necessarily admit of a corresponding inconsistent belief in the mind once adopted. The possibility of

self-induced deception is itself polemical, but that is unrelated to the topic at hand.

PERMANENCE AND DEGREES

Is it possible for a state of self-deception to become permanent? It seems so. One can imagine a belief that admits of such depth of ingression that to remove it would be to essentially raze the believer's entire cognitive framework. Similarly, it seems possible that a set of related albeit inconsistent beliefs could become so ingrained within one's noetic structure that it practically cannot be removed. Take, for example, Jeff. Jeff believes that the United States is in imminent danger of international nuclear war with China. Despite all the evidence to the contrary (namely, the fact that the United States is *not*, in fact, in imminent danger of transcontinental nuclear war), he nonetheless has deceitfully convinced himself of this danger insofar as to construct an underground nuclear bunker in his backyard, in which he now lives. He has, rather literally, built both his life and perceptual framework upon the foundation of this self-contradictory set of beliefs. It seems that no amount of reasonable discussion or solicitation of evidence will convince Jeff that he entertains a set of beliefs that are, at least partially, self-contradictory (the U.S. is in danger and that it certainly seems that the U.S. is not in danger). Therefore, it is evident that the permanence of self-contradictory beliefs hinges on their depth of ingression, or how many beliefs are supported by at least one member of the inconsistent pair.

Consequently, it follows that the vice of self-deception can also come in degrees. The man we discussed above is ostensibly more self-deceived than Tom, the student mentioned earlier who refuses to admit that he will most likely fail his next math quiz. This discrepancy can be attributed to the fact that the man who fears nuclear war is no longer open to a change of his beliefs, while Tom most likely still possesses the willpower to right his inconsistent set of beliefs. Where then do these two diverge? As noted above, self-deception begins to solidify when more beliefs are built on top of the aberrant pair. But in addition to this, self-deception arises when one permits or even mentally encourages further self-deception upon the act of self-examination (Strandberg 2015, 49). Jeff, upon self-examination, will most likely conclude that his mind admits of no contradictory beliefs. Rather, while he cannot dispose of the belief that almost all available evidence discourages his belief that the U.S. is in danger, he will attempt to wrongfully justify this belief by

asserting that all this evidence is, in fact, misguided. This process only serves to deepen the depth of ingression and promote further self-deception.

PREVENTION

Thus, it can be seen that honest self-examination is a crucial ally in the battle against self-deception. However, there is more to the evasion of self-deception than honest self-reflection. To this point Augustine weighs in: “[Charity] produces humility and forestalls the tendency to exaggerate our own goodness...According to this view, then, the remedy for self-deception lies not in a person’s own self-reflective capacities; it comes by way of a moral transformation made possible by charity” (Floyd 2004, 77). According to Augustine, Floyd explains, *charity* is the key to avoiding self-deception. True charity is constituted by a reordering of one’s desires (with the help of the Holy Spirit) to pursue the proper objects, which leads to a mitigation of self-interest and pride. And, as I discussed earlier, pride is the root of self-deception. Thus, reordering one’s loves through a significant and continually ongoing “moral transformation” increases wisdom and diminishes pride, and enables one to fairly examine their own beliefs. This practice can and should be abetted through the assistance of others, as outside sources will almost always exhibit a more charitable approach to one’s cognitive structure. Finally, we must seek to practice the virtue of a strong will—this enables us to courageously and authentically pursue true belief in any circumstance and avoid those vicious states of self-deception.

In the last analysis, self-deception seems to us a vice that always afflicts those around us (particularly those who exasperate us), but never one that affects our own mental life. And yet, ironically, to believe this notion is to deceive ourselves, as each one of us has been affected by a state of self-deception at some point. In this paper I discussed the possibility of self-deception and concluded that it is not only possible, but prevalent in noetic convention. I then defined self-deception as a habitual tendency to allow two or more self-contradictory beliefs to exist in one’s mind due to a sort of cognitive barrier. Moreover, I discussed its causal ancestry, characteristic thoughts and actions, related vices and virtues, as well as its permanence and degrees. Finally, I discussed strategies of avoidance: namely, the reordering of loves to humble us and permit the fair investigation of our mental state. Ultimately, we must remember that we deceive ourselves all the time—yet this is no cause for despair. Rather, we should acknowledge our vicious states, and

do everything in our power to diminish them through honest introspection and persistent authenticity.

REFERENCES

- Aquinas, Thomas. 1969. *Summa Theologiae*, edited by Thomas Gilby. Garden City: Image Books.
- Coleman, Mitchell Carl. 2007. "Contribution of Thomas Aquinas's Treatise on Temperance to the Contemporary Effort to Understand and Treat Addiction." MA diss., University of Iowa.
- Davidson, Donald. 2010. *Problems of Rationality*. Oxford: Clarendon Press.
- Davidson, Donald. 2013. *Essays on Actions and Events*. Oxford: Clarendon Press.
- Dupuy, Jean-Pierre. 1998. *Self-Deception and Paradoxes of Rationality*. Stanford: CSLI Publications.
- Elshof, Gregg Ten. 2009. *I Told Me so: Self-Deception and the Christian Life*. Grand Rapids: Eerdmans.
- Floyd, Shawn D. 2004. "How to Cure Self-Deception: An Augustinian Remedy." *Logos: A Journal of Catholic Thought and Culture* 7 (3): 60–86.
- Leeuwen, D. S. Neil Van. 2009. "Self-Deception Won't Make You Happy." *Social Theory and Practice* 35 (1): 107–132.
- LePore, Ernest, and Brian P. McLaughlin. 1992. *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. New York: Blackwell.
- McLaughlin, Brian P., and Amélie Oksenberg Rorty. 1988. *Perspectives on Self-Deception*. Berkeley: University of California Press.
- Mele, Alfred R., and Piers Rawling. 2010. *The Oxford Handbook of Rationality*. Oxford: Oxford University Press.
- Roberts, Robert Campbell, and W. Jay Wood. 2012. *Intellectual Virtues: an Essay in Regulative Epistemology*. Oxford: Clarendon Press.
- Strandberg, Hugo. 2015. *Self-Knowledge and Self-Deception*. London: Palgrave Macmillan.