

Editorial Manager(tm) for Erkenntnis
Manuscript Draft

Manuscript Number: ERKE417R2

Title: The Product of Self-Deception

Article Type: Original Research

Section/Category:

Keywords: self-deception; belief; avowal; avowed belief; paradox

Corresponding Author: D. S. Neil Van Leeuwen,

Corresponding Author's Institution:

First Author: D. S. Neil Van Leeuwen

Order of Authors: D. S. Neil Van Leeuwen

Manuscript Region of Origin:

Abstract: I raise the question of what cognitive attitude self-deception brings about. That is: what is the product of self-deception? Robert Audi and Georges Rey have argued that self-deception does not bring about belief in the usual sense, but rather "avowal" or "avowed belief." That means a tendency to affirm verbally both privately and publicly, but lacking normal belief-like connections to non-verbal actions. I contest their view by discussing cases in which the product of self-deception is implicated in action in a way that exemplifies the motivational role of belief. Furthermore, by applying independent criteria of what it is for a mental state to be a belief, I defend the more intuitive view that being self-deceived that *p* entails believing that *p*. Beliefs (i) are the default for action relative to other cognitive attitudes (such as imagining and hypothesis) and (ii) have cognitive governance over the other cognitive attitudes. I explicate these two relations and argue that they obtain for the product of self-deception.

Response to Reviewers: Dear Editor and Referees,

Many thanks for this last round of comments. I have attempted to make appropriate revisions in response to the comments received. Relevant changes to the last revision include the following. (1) I have fleshed out the abstract to suggest better my positive view on beliefs. (2) I have slightly changed the numbering scheme to avoid the awkwardness indicated by Reviewer #2. (3) I had a proofreader go over the text, as suggested. (4) I have made clearer that 1.2 is not meant to be a comprehensive literature review, but rather a sketch of exemplars to motivate the present investigation. (5) I modified my response to Audi's case of Jan the teenager who appears to contemplate suicide (p. 12-13). (6) I consider in endnote 22 the objection that Audi or Rey could make that the cases I discuss are not cases of self-deception. In addition, there have been a few other changes of style to improve clarity. I have also added acknowledgements.

All the best,
The Author

The Product of Self-Deception

“It does not appear to me that S’s being in self-deception with respect to p and unconsciously believing not- p entails his believing—albeit consciously—that p . All my view requires regarding S’s positive attitude toward p is that S be disposed sincerely to avow it.” –Robert Audi (*Erkenntnis* 18, 1982)

Abstract: I raise the question of what cognitive attitude self-deception brings about. That is: what is the product of self-deception? Robert Audi and Georges Rey have argued that self-deception does not bring about belief in the usual sense, but rather “avowal” or “avowed belief.” That means a tendency to affirm verbally both privately and publicly, but lacking normal belief-like connections to non-verbal actions. I contest their view by discussing cases in which the product of self-deception is implicated in action in a way that exemplifies the motivational role of belief. Furthermore, by applying independent criteria of what it is for a mental state to be a belief, I defend the more intuitive view that being self-deceived that p entails believing that p . Beliefs (i) are the default for action relative to other cognitive attitudes (such as imagining and hypothesis) and (ii) have cognitive governance over the other cognitive attitudes. I explicate these two relations and argue that they obtain for the product of self-deception.

1 Introduction

Suppose your friend’s son is underperforming in school, doesn’t read a lot, and doesn’t get the jokes that most kids his age get. Suppose also your friend insists that his son is “very smart” and in fact “smarter than most other kids his age.”

Knowing your friend almost never lies and certainly wouldn’t to you, and knowing him to be no fool in general, you see that he’s deceiving himself. But now, since you’re a philosopher, the usual puzzle arises in your mind. You think that at one level he must *know* that his son is *unintelligent*. But he keeps telling you with a sincere voice that his son *is* intelligent, so he must *believe* that his son is intelligent. If knowledge entails belief—as we often say it does—then your friend, in order to be deceiving himself, must both believe that his son is intelligent and believe that his son is not intelligent. More formally, he believes that p and believes that $\sim p$. But this seems psychologically absurd. So we have arrived at the well-known paradox of self-deception: start with an ordinary description of everyday mental life—a person is deceiving himself—and on quite straightforward analysis of the terms involved in the description see that it entails absurdity.

1.1 Five Possible Approaches to the Paradox

What are some strategies for solving the problem? Five are prominent:

1. *Deny the phenomenon.* Some say self-deception doesn't exist. The absurdity derivable from semantic analysis of the term "self-deception" shows that no real phenomenon can answer to it. The advantage of this strategy is that you can dispel the paradox without much mess. The danger is that you might be denying a real phenomenon.

2. *Go Freudian.* The next strategy is to divide the mind into distinct centers of agency, with a subsystem intentionally deceiving the main system. On Freudian views, the separate parts *act as distinct agents* with *their own* desires, goals, beliefs, and intentions. The advantage of this approach is that one can describe self-deception like other-deception, which lacks paradox. The disadvantage is that it may make the ordinary self-deceiver out to be too much like a split personality case, in which the agential incoherence seems to far outstrip that of the ordinary self-deceiver.¹

3. *Hide one of the beliefs somewhere else.* One needn't posit separate centers of *agency* within an individual in order to say that the mind has different parts, the distinctness of which may prevent beliefs within the separate parts from interacting as they would in the same part. One may execute this strategy by distinguishing between conscious beliefs and unconscious beliefs, but there are other possible ways that beliefs might be compartmentalized. The main advantage of this strategy is that it captures the idea that two beliefs are involved in self-deception *without* positing distinct, homuncular centers of agency. The disadvantage is that it can make self-deception appear like a simple error in processing. Where's the *deception*?

4. *The avowal view.* Perhaps your friend doesn't *really* believe that his son is especially intelligent. He's merely disposed to talk that way. He has what one might call an *avowed belief*, a state that explains how a person is disposed to speak sincerely², but which lacks other properties of beliefs, like deep connections to (non-verbal) action. Let's call the attitude that results from self-deception the *product of self-deception*. So in your friend's case the product has the content *my son is highly intelligent*. The avowal view holds that the product is an avowal. The advantage of this view is that it escapes the paradox apparent in positing two contradictory beliefs. The

disadvantage—since the product isn't a belief—is that it seems to rule out using self-deception to explain other, non-verbal actions. (As the opening quotation indicates, Audi is a proponent of the avowal view—more on his version of it below.)

5. *The one-belief view.* The final strategy is to say that the self-deceiver really only has one belief, namely, she believes the proposition she's deceived about; she doesn't believe the proposition that the evidence favors. I'll call this second proposition the *doxastic alternative*. So your friend—if this is right—doesn't know or believe his son is unintelligent, although in some sense he should; he only believes his son is intelligent. This approach allows us to employ self-deception to explain actions beyond merely verbal ones. But there are two disadvantages. First, this view may miss the epistemic tension one might think is inherent in self-deception. Second, positing belief in the doxastic alternative seems needed to explain behaviors characteristic of self-deceivers, like avoiding situations that would force them to confront the uncomfortable truth.³

1.2 The State of Scholarship on the Metaphysics of Self-Deception

Puzzled about how to describe the state of your friend, you consider two basic questions. Is the *product of self-deception* a belief? Does the self-deceiver believe the *doxastic alternative*? The self-deception literature reveals two facts. First, *there is no consensus* on how to answer these questions or what strategy to take. Second, positions that appear to have died can come back. (The literature on self-deception is sizable, so in this section I only attempt to pick out examples of the various positions in order to give the reader a sense of the scholarship.)

There is fairly general consensus in the literature that self-deception exists. So we might think that *deny the phenomenon* went out with the days of hard-nosed, literalistic ordinary language conceptual analysis. Steffen Borge, however, has argued as recently as 2003 that the very notion of self-deception is incoherent.

Those who affirm the existence of self-deception differ on whether the self-deceiver believes the doxastic alternative. Davidson (1982, 1985, 1998), as a Freudian, affirms that the doxastic alternative is believed. McLaughlin (1988), on the other hand, is a non-Freudian who

chooses the *hide one of the beliefs somewhere else* strategy and hence still affirms belief in the doxastic alternative. The avowal theorists, including Audi (1982, 1988) and Rey (1988, forthcoming), affirm there is a belief in the doxastic alternative, which pressures them (to avoid the paradox) into holding that the product is not a belief. But there are those who hold that it is not necessary—and not likely—that the self-deceiver believe the doxastic alternative, including Talbott (1995) and Mele (2001).⁴

Turning to the second question, you find that the majority view is that the product of self-deception is a belief. But the position of the vocal minority of those who adopt *the avowal view*, Audi and Rey, seems never to have been answered. Furthermore, what looks like a version of the avowal view appears in Funkhouser (2005).⁵

Although the literature is a mess, there are strands of consensus. First, it is widely agreed that some motivational attitude is constitutively involved in causing self-deception; I shall call this the *deceptive element*.⁶ Second, it's uncontroversial that the self-deceiver has to have some sort of access to information that would justify believing the doxastic alternative.⁷ Third, it seems agreed that the product of self-deception is some cognitive attitude, where a *cognitive* attitude is one that can be evaluated as true or false—as opposed to conative attitudes, like desires.⁸

So there is lack of consensus about specific details of self-deception *amid* vague consensus about some of its critical components. The vague consensus indicates that we philosophers are at least attuned to the same phenomenon (*mirabile dictu*). To improve the dialectic, however, we must incorporate into the study of self-deception *independent* characterizations of the states that may be involved. Importantly, we will have to be more explicit about what *belief* is.

1.3 The Question to be Addressed: What is the Product of Self-Deception?

The goal of this paper is specific, as I cannot here address all of the problems raised by lack of consensus in the literature. I will argue that the *product of self-deception* is a *belief*; it is

not merely an avowal.⁹ I'll call this view the *belief view*, encompassing all views that hold the product is a belief; the opposed view is the *avowal view*.¹⁰

Here's the structure of my argument. In section 2, I present a taxonomy of different forms of self-deception based on the deceptive/motivational element; this taxonomy should help minimize cross-talk and provide a framework for future hypothesis formulation. The taxonomy also presents examples that will serve as a basis for further analysis. In section 3, I present the analyses and motivations of those who hold the avowal view, focusing on the work of Robert Audi and Georges Rey. In section 4, I argue that the avowal view lacks motivation independent of being an attempt to resolve the paradox. I then show in section 5 why the avowal view fails, arguing that it fails to make sense of the connection that self-deception has to non-verbal behavior and that it seems to confuse self-deception with a phenomenon that might more appropriately be called *flakiness*. In section 6, I develop independent criteria of what it is to be a belief and apply them to the product of self-deception. Section 7 concludes with a summary and a note on moral psychology.

2 A Taxonomy of Forms of Self-Deception

Let's hold fixed that the deceptive element is some form of motivation. For convenience I'll refer to this as a desire, although I wish to leave open whether other, emotional forms of motivation can be implicated in causing self-deception. Examining cases will reveal that, in self-deception that *p*, desires with a variety of *p*-related contents can be the deceptive element.¹¹

2.1 Wishful Self-Deception

Let's consider a business owner. Her business has been successful until a year ago, but has been running at a bad loss and shows no sign of getting better. In fact—and this is the intuitive test for self-deception—this owner has information such that, if she were to see it in relation to someone else's business, she would believe that *that* business will surely fail. Nevertheless, she wants badly for her business to succeed and insists that it will turn around. She

has the opportunity to sell off her current assets for a considerable sum. Instead, she takes out loans against those assets such that everything will go to the bank if the business fails.

In this case, the deceptive element is the desire *that this business will become successful again*. The product is the cognitive attitude with the content *that this business will become successful again*. Thus, in one type of self-deception, the content relation is identity. I call this type *wishful self-deception*, since it's continuous with wishful thinking.

2.2 Willful Self-Deception

Let's add more detail to the case from the beginning of the paper. Your friend holds that anyone who doesn't believe his child is excellent in every respect isn't a good father. Thus, he wants to *believe* his son is intelligent. He may also want it *to be the case* that his son is intelligent, but let's suppose it's his guilt-ridden desire *to have the belief* that's driving the self-deception.

Thus, the second kind of relation between content of deceptive element and content of product is higher-order: the content of the deceptive element desire is *that the agent believe* what ultimately becomes the content of the product of self-deception.¹²

2.3 Goal-Driven Self-Deception

In 1980, Air Florida Flight 90 crashed. As take-off was underway in icy conditions (but still reversible), the co-pilot warned the pilot about problematic readings on the instrument panel. The pilot, however, allowed take-off. The pilot, we may reasonably infer, was self-deceived. The product of his self-deception has the content *that the plane is safe enough to fly*. But what is the deceptive element? We might think that he simply wants it to be the case that the plane is safe. But really the motivating force behind the self-deception seems to be the *goal* he has to get going.¹³

Thus, in *goal-driven self-deception*, the deceptive element is a goal, and the content of the product of self-deception is that a means to achieving that goal is satisfied. In order for the pilot's goal of taking off quickly to be satisfied (deceptive element), the plane has to be safe enough to fly (content of product).¹⁴

2.4 Dreadful Self-Deception

People can also self-deceptively believe what they want *not* to be the case. I call this *dreadful self-deception*, because in it the self-deceiver believes what she in some sense dreads. Eagerly desiring that her boyfriend *not* be cheating on her, a person taken by dreadful self-deception believes that he is, even though there is plenty of evidence that he's faithful. Thus, in dreadful self-deception, the content of the motivational element is the negation of the content of the product of self-deception.¹⁵

2.5 What Holds these Different Types of Self-Deception under one Concept?

There are synchronic and diachronic commonalities that join these different types of self-deception as a kind. Synchronically, they are all characterized, as noted, by the holding of a cognitive attitude contrary to the agent's evidence and epistemic norms, where this holding is under the influence of a desire or other motivational component. Diachronically, each kind of self-deception involves the influence of that motivational component on *attention* in a way that produces and supports the cognitive attitude or product of self-deception. The attention is pushed by aspects of the motivational element *away* from the greater evidence that supports belief in the doxastic alternative and onto the scantier evidence that supports the product of the self-deception. What makes this a process of *self-deception* is that the agent feels a rational pull to attend to evidence she's not attending to (which would dictate believing the doxastic alternative if properly considered), but she motivatedly ignores that evidence nonetheless. This is not so analogous to straightforward interpersonal lying; it is more analogous to withholding pertinent information and emphasizing of misleading information in a way that leaves a false impression, which is still a form of deception.¹⁶

The motivated modulation of attention figures in the production of the product of self-deception in all of the types of self-deception mentioned above. The business owner's holding that the business will turn around is supported by the few signs of promise and the few good days of sales that she focuses on. The father who holds that his son is highly intelligent fixates on the

few clever things his son has said. The pilot whose plane crashed focused on meter readings that looked promising. The young woman in dreadful self-deception that her boyfriend is unfaithful focuses on the friendliness of the conversation he had with her friend.

3 The Avowal View: Audi and Rey

The self-deceived business woman represents her business as about to succeed. The Air Florida pilot represented the plane as sufficiently safe to fly. In what sort of mental states are these representations embedded? Audi (1982, 1988, and 1997) and Rey (1988) would, in order to resolve the paradox of self-deception, say they're avowals or avowed beliefs.

How does this solve the paradox? An avowal, for both Audi and Rey, is—roughly—a disposition to affirm a proposition to oneself and others with “sincerity,” but which lacks deep connections to action. Audi, but not Rey, adds that this avowal is conscious.¹⁷ Thus, if the self-deceived agent merely avows that p in self-deception that p , we needn't attribute to the agent both the belief that p and the belief that $\sim p$. When an agent is self-deceived that p , both Audi and Rey hold that the agent has a belief (which Rey calls a “central belief”) that $\sim p$.

Divorcing avowals from connection to action is a critical move for both Audi and Rey, for otherwise it is completely unclear in what respect an avowal is actually supposed to be distinct from a genuine belief. To define what avowal is Rey (1988, p. 278) says: “To a first approximation, one could say that, *ceteris paribus*, a person avowedly believes that p if she would sincerely and decidedly assert p if asked.” Avowed beliefs, for Rey, are those we attribute on the basis of verbal behavior. Central beliefs are those we attribute on the basis of actions. Along the same lines, Audi (1982, p. 138, quoted earlier) adds: “It does not appear to me that S's being in self-deception with respect to p . . . entails his believing—albeit consciously—that p . All my view requires regarding S's positive attitude toward p is that S be disposed sincerely to avow it.”

It is clear in the case of Audi and strongly suggested by Rey that avowals are not causes of non-verbal action in the way that beliefs are, so the distinction between avowals and beliefs is not empty. First Audi (1982, p.139), in considering whether one should attribute a genuine belief

that p in a certain case of self-deception that p , says: “her overall *behavior* will not support the attribution of this belief” (my emphasis). Second Rey (1988, p. 264), in his initial characterization of a case of self-deception, says: “In self-deception people seem straightforwardly to fail to *act* in accordance with the attitudes they sincerely avow: sincerely claiming her lover to be intelligent, someone nevertheless seldom takes his view seriously . . . she ‘conceals from herself’ her real opinion” (my emphasis). Again, the avowal, the product of self-deception, is divorced from action.

What else would Audi and Rey say about the business woman and the pilot? Audi (1982, p. 137) holds that:

S is in *self-deception with respect to p* if and only if:

- (1) S unconsciously knows that not- p (or has reason to believe, and unconsciously and truly believes, not- p)
- (2) S sincerely avows, or is disposed to avow sincerely, that p ; and
- (3) S has at least one want which explains in part both why the belief that not- p is unconscious and why S is disposed to dis-avow a belief that not- p , and to avow p , even when presented with what he sees as evidence against it.

Audi’s first clause commits him to the view that the business woman unconsciously *knows* (or with reason truly believes) that her business will not succeed (this is the “not- p ” or doxastic alternative). Likewise, Audi would say the pilot unconsciously knows the plane is not safe. I think Audi’s view that the self-deceiver has knowledge that the doxastic alternative is true is a large part of what pushes him in the direction of the avowal view. Knowledge that the doxastic alternative is true entails belief that it is, which seems incompatible with belief in its contrary; hence the need to posit some *other* kind of mental state as the product of self-deception, i.e., avowal, which comes out in Audi’s second clause.

But we should critically evaluate Audi’s knowledge commitment. Independently of knowing that one of these characters has fallen into self-deception, do we have good reason to attribute *knowledge* to the business woman or the pilot—conscious or unconscious—that runs contrary to their respective self-deceptions? There is an often represented line of reasoning that claims that positing knowledge or belief with the doxastic alternative as content is *needed* to

explain certain self-deceptive behaviors.¹⁸ For example, perhaps the businesswoman avoids reviewing her balance sheets; some would argue this indicates she knows or at least believes what they would tell her, i.e., that her business is failing. In short, if a person is self-deceived that p , she must believe that $\sim p$ (at some level) in order to have proper motivation to avoid the evidence as she does. To my mind, however, this type of argument suffers from failure to consider the full range of possible cognitive states that could underlie the mooted avoidance behavior; belief is one of many. An uncomfortable *suspicion* that $\sim p$ (the doxastic alternative), for example, could be enough to motivate one not to ask the questions that would undo her self-deception. What matters is that the agent has evidence favoring $\sim p$, which is not adequately attended to due to the deceptive element, and that this evidence creates tension with the product of self-deception. Saying that the doxastic alternative is encoded in *knowledge* or *belief* results in a badly misleading description of the mental state of the self-deceiver.¹⁹ The sad thing about self-deception, I think, is that the people *lack* knowledge about themselves or the world. The business woman *doesn't* know where her business is headed, although she perhaps *should*. The pilot doesn't know what he should about the safety of the plane, despite what may be nagging suspicions that he should reconsider the matter. So Audi's first clause, the knowledge clause, is problematic; rejecting it, we also lose the pressure toward the second clause, the avowal clause. Audi would hold that the business woman doesn't actually *believe* that her business will turn around. I find this dubious.

Rey (1988, p. 281) writes: "The following should serve as a first approximation of self-deception: an agent centrally holds some attitude that p and knows that she does; but, preferring not to avow it, she prevents herself from doing so, sometimes bringing herself to avow not- p instead." (Rey never goes beyond a first approximation.) This is an analysis of self-deception that $\sim p$. Of course, we could just as well flip the p and $\sim p$.

Rey's first clause is unclear. Rey says the agent centrally holds "some attitude that p ." But the range of possible attitudes isn't specified. Could the agent have a central *hope* or *desire*

that p and still be self-deceived that $\sim p$ just because she avows $\sim p$? This would not be self-deception; I could desire that p and know I do, but avowing $\sim p$ wouldn't make me a self-deceiver. It could mean I'm realistic. Maybe Rey really means that the agent centrally *believes* p and knows she does, but avows $\sim p$. But at what level is the "knows" supposed to operate here, avowed or central? It must be central knowledge, for otherwise we would be stuck with the paradox again at the avowal level. So let's assume Rey means the attitude in question to be belief and the knowledge to be central. In any case, it's odd to attribute so much knowledge to a self-deceiver.

There's also something problematic about the second clause in Rey's definition. Rey talks of the agent as *preferring not to avow* what she centrally believes. The problem is that preferring *not* to avow something does *not* in general explain why someone would avow the contrary. Given my ignorance of where the capital of Florida is, my preference *not* to avow that Miami is the capital of Florida doesn't predict (or explain) that I'll avow that Miami *isn't* the capital Florida. The truth is I don't know either way. Perhaps it would be charitable to attribute to Rey the view that the real preference that is constitutively operative in self-deception is not the negative one *not* to avow, but rather the positive one *to* avow what ultimately does get avowed.

How do the cases work out on Rey's definition? I'll continue focusing on the two to which I applied Audi's definition. First, the business woman centrally believes that her business will fail and knows that she centrally believes this. But she prefers to avow that her business will become successful again, where the avowal (on Rey's account) is something that is said both to herself and others with *sincerity*. Rey's appeal to the sincerity of the avowal is puzzling, given the knowledge that Rey claims the agent has of herself believing the contrary. Furthermore, it seems that not attributing to the business woman the genuine belief (or "central belief") that her business will go as she avows is unmotivated aside from an attempt to resolve an apparent paradox. What about the pilot? Rey is committed to saying that he centrally believes that the plane is unsafe and that he knows that he does. But, on Rey's story, the pilot prefers to avow that the plane is safe, so he does avow this.

How does the belief view, by way of contrast, deal with the cases? The business woman *believes* that her business will succeed; the pilot *believed* that the plane was sufficiently safe to fly. The cause of the epistemically illegitimate belief (to go with the common view) is a desire, or something like it—perhaps a content-bearing, motivating emotion.²⁰ In addition, the belief view is open as to whether the self-deceived agent actually believes the doxastic alternative. Does the business woman at some level *believe* that her business will fail? Davidson would say that she does; Mele would say that she doesn't. I'm inclined to say she doesn't, since one can capture the epistemic tension in self-deception by saying the businesswoman has compelling *evidence* of impending failure to which she motivatedly avoids attending. This evidence, despite her ignoring it, may surface as an uncomfortable suspicion. But this is not a matter to be decided here, since it's the *product* of self-deception that's at issue.

We've seen how the views on offer deal with the cases of self-deception I presented. Before attacking it directly, let's see whether the avowal view has independent motivation.

4 Independent Motivation for the Avowal View?

Is the avowal view just a philosopher's trick to avoid paradox? Both Audi and Rey attempt to supply some independent motivation for the avowal view. I argue they're unsuccessful.

First, Audi (1988, pp. 94-95) presents the case of Jan. Jan is a teenager who engages in a suicide "attempt" by taking only six aspirin, leaving the open bottle out for her parents to see. Then she tells her friends and parents about the "attempt." Her real motivation is to get attention. So she's self-deceived that she's "seriously contemplating suicide."

What's Audi's argument that she doesn't really *believe* (and rather merely avows) that she's seriously contemplating suicide? "She does not quite believe what she says, however. She is too aware of her own behavior and feelings for that; and this is why we do not expect the full range of behavior one would expect from genuine belief, including planning that presupposes that she will die . . ." (1988, p. 95) The idea is that, if she really *believed* she's seriously

contemplating suicide, she would plan and behave like she's going to die. But she doesn't so behave, so she must not actually believe—she merely avows. Or so the argument goes.

The argument, however, is flawed. *Believing* you're seriously contemplating suicide is not necessarily supposed to get you to act and plan like you're going to die (although it may do this); it's *actually being suicidal* that does this. Jan may believe she's seriously contemplating suicide, but if this belief is *wrong*—as we'd expect if she's self-deceived—then she might not *act* like she *is* suicidal. Audi seems to confuse the expected effects of *believing* one is suicidal with the effects of *being* suicidal. Jan's not fully behaving like she's suicidal doesn't entail that she doesn't believe she is seriously contemplating suicide. Furthermore, if she does believe she's seriously contemplating suicide, the effects this belief will have on action will be dependent upon what her relevant desires are. Jan's not planning to die may just as well be indicative of a lack of desire for such plans to be in place as it is of a lack of belief that she's seriously contemplating suicide. So Audi's example is inconclusive.

If we grant Audi the distinction between avowals and beliefs, it's true that I haven't yet given reason for rejecting the view that Jan merely avows she's seriously contemplating suicide. But the point is that Audi's argument for seeing the case this way as opposed to adopting a belief view doesn't work; his appeal to her behavior doesn't establish lack of belief. So all we're left with for motivation for his avowal view is the initial paradox; there's no independent motivation.

Does Rey offer independent motivation for making the product of self-deception an avowal? Rey devotes more attention to arguing that there is a distinction between central and avowed states than to justifying the view that the divergence of the two is characteristically what's involved in self-deception. He doesn't present his definition of self-deception that makes appeal to avowed belief until section 6 of his article, where his defense is not extensive. We can accept his distinction, but still maintain that the product of self-deception is a genuine belief. There's one obstacle to saying this, and that's the case that Rey describes in the beginning of the article. Here it is again: a woman is in a state of self-deception about the intelligence of her lover;

she says she thinks he's intelligent but never takes his opinions seriously, etc. Rey's thought must be: she's self-deceived about her lover's being intelligent, but she doesn't act like she believes it, so the product of self-deception must not be a genuine belief. But the case is under-described. Does the woman have evidence her lover is unintelligent? If so, what does she do with it? What sorts of utterances is she dismissive about? Do considerations aside from his verbal behavior support the view that he *is* intelligent? My inclination is to say that if this really is a case of self-deception, then she does believe he's intelligent. If she doesn't take his views seriously that can be explained as the product of her independent assessments of each view. Nothing Rey says rules out this interpretation. And if she doesn't believe, then she's not really self-*deceived* that he's intelligent. She merely likes having the thought in mind.²¹ But, again, the case is under-described.

I conclude that neither Audi nor Rey gives motivation independent of the paradox itself for thinking that the product of self-deception is avowal rather than belief.

5 Why the Avowal View won't Work

Consider again the business woman who's self-deceived that her business will become successful again. Should we say she merely *avows* this? Or should we say that she *believes* this? If she doesn't actually believe her business will turn around, how does it come about that—at the very time it would be best to cut losses—she takes out a loan to save the business? Recall that avowals for both Audi and Rey are distinguished from proper beliefs by lacking the connection to action we normally understand beliefs to have. If they were so connected, it wouldn't be clear what an avowal as opposed to a belief is supposed to be. If you say the avowal in the business woman case really does play a causal role in action, then what you're talking about is a belief.

If the product of self-deception is a belief, we can explain why the business woman takes out the loan. She wants the business to turn around and believes that the loan will enable her to turn it around. Behind this latter belief is the general, action-guiding belief that the business will succeed again. What does the proponent of avowals say here? Audi and Rey can posit a plethora of desires to save face in business regardless of success or failure, but if either of them is right,

the business woman essentially knows her business will fail. Assuming she also wants to minimize loss, it's hard to imagine that she'd take out the loan with knowledge of impending failure. But she does, and the reason why is that she *believes* the business will turn around.

Can the same point be made about the other cases considered? I would argue that the case of the pilot very clearly shows a product of self-deception that is a belief: *he allowed the plane to take off*. It's hard to imagine a pilot with central knowledge that a plane isn't safe (and only an avowal that it is) allowing take off. A much better explanation is that he believed the plane was safe enough to fly. The other cases—the father self-deceived about his son's intelligence and the girlfriend self-deceived about her boyfriend—can't distinguish between the avowal and belief views, since the evidence of the product is largely verbal behavior. But the pilot and the business woman do establish that for a significant class of cases of self-deception the product is belief.²²

There is another salient argument against the avowal view. We can contrast the self-deceived business woman with another person we might call merely flaky or superficial. Imagine a business woman whose business is failing in like fashion to that of the self-deceived business woman. This one, however, doesn't take out the loans or attempt to save the business in any way, despite insisting that she can turn the business around. The business woman who doesn't take action is just being flakey, talking one way and acting in a way that indicates contrary beliefs. How do we describe her psychology? We can best put it in Rey's terms: *she centrally believes that her business won't succeed and knows that she does; but, preferring not to avow it, she prevents herself from doing so, bringing herself to avow that it will succeed instead*. So Rey has a good description for something, but it's not self-deception; it's *flakiness*.

Thus there are two reasons for holding that the product of self-deception is a belief. First, it allows us to make sense of the connections self-deception has to action. Second, it allows us to distinguish self-deception from flakiness. The self-deceiver believes; the flake merely avows.

6 An Independent Argument for the Belief View

Beliefs are the most basic attitude in the human cognitive economy. Having bodily movements based on beliefs is also the difference between action and reflex. In the next two subsections (5.1 and 5.2), I spell out some essential properties of beliefs and argue that these properties obtain for the product of self-deception.

6.1 What are Beliefs?

Beliefs can be characterized by the relations that obtain between them and other cognitive attitudes. I will look at two pertinent types of relation here: cognitive and practical.

6.1.1 The Cognitive Role of Beliefs

On the cognitive side, beliefs (i) are conditions of possibility for the existence of other cognitive attitudes and (ii) play systematic roles in determining their contents.

This can be seen by looking at hypothesis. An hypothesis is a mooted explanation for why a situation is a certain way. Thus, in order to have hypotheses at all, one must have *beliefs* about the situations that call for explanation; at the very least, one believes that the situation exists. In addition, beliefs determine the acceptance or rejection of hypotheses. If I *believe* evidence that conforms to an hypothesis, then I accept it; if I believe evidence that contradicts an hypothesis, I reject it.²³ Furthermore, other beliefs one has about a situation determine the content of what we may call the *forward projection* of an hypothesis about that situation; the forward projection is the structure of anticipations about what the world will be like if the hypothesis is true. For example, if I hypothesize that the dog has gotten into the cupboard and believe that he likes Wheaties, then the forward projection of my hypothesis will include the anticipation that the Wheaties will have been gotten into; if I believe the dog likes crackers, then my forward projection will include the anticipation that the crackers will have been gotten into. I summarize the relations beliefs bear to hypotheses by saying that beliefs *govern* hypotheses, i.e., provide grounds for their existence and systematically determine the content of the forward projections.

Beliefs also govern imagining. Suppose I were to ask you to imagine an apple. In order for you to accomplish this, you must have *beliefs* about what apples are like; otherwise you might

as well say you *can't* imagine an apple. An instance of imagining can be divided into the *ground*, which is a tokening in thought of the concept of whatever it is you're trying to imagine, and the *scene*, which is the mixed sensory and conceptual reconstruction in thought of a representation of something that satisfies the ground. In the apple case, the ground involves the *apple* concept, while the scene consists of your visualization, other thoughts, and olfactorizations of the apple you are imagining. Given these distinctions, we are in a position to note that *beliefs* determine the production of essential and characteristic features of the scene on the basis of the ground. It is because you *believe* that apples have greenish-white flesh that your visualization of the flesh of the apple you imagine is greenish white, and so on. Thus, as in the case hypothesis, belief has cognitive governance over imagining. Similar governing relations hold between beliefs and other cognitive attitudes, such as acceptance in a context.

6.1.2 The Practical Role of Beliefs

Turning to the practical side, we have already noted that beliefs play a role in the constitution of action. Beliefs, in general, represent for us the way the world is, such that, if the world really is that way, the actions we choose will satisfy our desires. It is clear that the product of self-deception can act as input into practical reason in the way that other beliefs do, as the case of the businesswoman shows. But the argument is not done here, since other cognitive attitudes may also figure in the constitution of action. Wanting milk, for example, I may perform the action of going to the refrigerator simply on the *hypothesis* that there is milk in it. *Imagining* a hurdle before me, I may jump. So how is it that beliefs *differ* from the other cognitive attitudes in the constitution of action?²⁴

To answer this question, we can take our lead from Bratman (1992), who examines the difference between belief and acceptance in a context:

An agent's beliefs provide the *default cognitive background* for further deliberation and planning. . . . this cognitive background is, in the sense explained, context independent. But practical reasoning admits adjustments to this default cognitive background, adjustments in what one takes for granted in the specific practical context. . . . To be

accepted in a context is to be taken as given in the adjusted cognitive background for that context. (pp. 10-11, Bratman's emphasis)

If we extend this idea from acceptance in a context to other non-belief cognitive attitude, we can see the following point in this passage: non-belief cognitive attitudes require specific contexts in order to function as the background of deliberation for the constitution of action. When an imagining prompts action, it is *in the context of a game of make-believe*; when an hypothesis prompts action, it is *in the context of an investigation*. Beliefs, however, are the default for the constitution of action; their role in the causation of action is not limited to specific types of situation. For example, my belief that cats can scratch determines action: when I'm near a cat; when I'm telling someone else about cats; when I'm thinking about whether to put one in the same room with a dog; if I should ever think about becoming a veterinarian; when I am choosing whether to trim the cat's claws myself or have someone else do it; and so on. The action-causing role of beliefs is, as Bratman puts it, context independent.

Furthermore, when an agent acts on another cognitive attitude because of being in a certain context, it is because she *believes* she is in that context. If I jump an imagined hurdle, it is because I believe I am playing make-believe. If I act on the assumption that costs will be high (even if I don't believe it), I am acting on an acceptance in a context, *believing* that context to be asymmetry of costs of errors.

To summarize, belief's being the default for action relative to the other cognitive attitudes encompasses two aspects. First, the role of beliefs in the causation of action is context independent. Second, when an agent acts on another cognitive attitude because of being in a certain context, this is possible because she believes herself to be in that context. Action on the basis of other cognitive attitudes tends to occur only when beliefs on the subject matter of those other attitudes are lacking, or in a specific kind of context, where beliefs represent context.²⁵

6.2 Applying this Framework to the Product of Self-Deception

If what I have said about the relations between beliefs and other cognitive attitudes is correct, we can determine whether the product of self-deception is a belief by seeing whether the relations obtain. In other words, we must ask: does the product of self-deception govern other cognitive attitudes in the manner of belief? is the product of self-deception the default for action relative to the other cognitive attitudes? I argue here the affirmative to these questions.

We can ask, for example: how does the product of self-deception interact with hypotheses? Suppose that the father self-deceived about his son's intelligence hypothesizes that his son's next set of grades will be reflective of the son's intellectual ability. It seems likely that the forward projection of this hypothesis will include the anticipation that the grades will be strong, which means that the product of self-deception (which represents the son as being intelligent) is governing contents of the forward projection of the hypothesis in the way that a belief would. Furthermore, suppose that the grades turn out to be poor. If the product of self-deception leads the agent to reject the hypothesis that the grades would be reflective of the son's intelligence—which seems likely—then that product is clearly playing the governing role of beliefs over hypothesis. Such examples can be multiplied.

Does the governing relation obtain between the product of self-deception and acceptance in a context? Suppose a contractor has a policy of reasoning as if the costs of his projects will always be in the upper end of his projected range whenever a project involves some risk, doing so in order to avoid disastrous loss. In other words, he accepts that costs will be high in the context of risk. But suppose he is self-deceived about how much his friend will charge him for wood, self-deceptively holding that—as he is a friend—he will charge less. Will the product of this self-deception affect the acceptance in a context about the high end of the cost just as a *belief* about the “friend” would? That is, would the product of self-deception lower the estimated high end that is accepted in the context? If so—and I think *yes* is the answer—this is further evidence that the product of self-deception is a belief.²⁶

What about the practical relation between the product of self-deception and the other cognitive attitudes? Is the product of self-deception the default for action relative to the others as beliefs are? Let's return to the case of the self-deceived business woman. It is clear that her action of taking out loans to keep the business afloat is driven to a great extent by her self-deception. But is the cognitive attitude underlying this action context sensitive, i.e., only psychomechanically effective in certain kinds of context? It is more likely that its influence on action pervades many contexts, such as speech, planning various endeavors, taking out loans, and considering whether to quit. If this is the case, then the product of her self-deception forms the default for action relative to the other, context-sensitive cognitive attitudes.

These considerations give strong reason to hold that the product of self-deception is a belief. Let me summarize the main points in this argument. First, the product of self-deception can play the role of belief in the causation of action. Second, the product of self-deception seems to cognitively govern the other cognitive attitudes in the manner of belief. Third, the product of self-deception has the role of being the default for action relative to the other cognitive attitudes.

In the Introduction, I outlined five strategies for handling the paradox of self-deception: *deny the phenomenon*, *go Freudian*, *hide one of the beliefs somewhere else*, *the avowal view*, and *the one-belief view*. If self-deception exists and if the argument of this paper is correct, then the correct solution to the paradox will come from one of the three strategies that hold that the product is a belief. I have not resolved which strategy it is, as this was not my purpose (although it may be apparent that I prefer a form of one-belief view). I have, however, provided a model of how further questions in the dispute should be addressed. For example, to figure out whether or not the doxastic alternative is believed, one should take a broader perspective and apply a robust independent characterization of what it means to believe something. Only by adopting a broader perspective in this fashion can we escape the fruitless question-begging that has characterized much argumentation about the nature of self-deception. Self-deception is interesting *because*, among other reasons, it forces us to ask what belief is.

7 Reflections on Self-Deception and Moral Psychology

Let's review the dialectic. We saw early on that there is very little consensus in the literature about how beliefs figure into states of self-deception, although (or because) there are (at least) five strategies on offer for addressing this question. The goal of this paper was to bring more clarity to the discussion of what self-deception is by focusing on one component of self-deception: the product. We taxonomized different forms of self-deception and came up with examples. With these examples as a basis, we compared the view that the product of self-deception is an avowal with the view that it is a belief. Motivation for the avowal view was lacking. Motivation for the belief view was found by applying independent criteria for belief to cases of self-deception. Those criteria were (1) that beliefs cognitively govern the other cognitive attitudes and (2) belief is the default for action relative to the other cognitive attitudes.

The chief problem with the avowal view is that it fails to make sense of the connections that the product of self-deception has to action. These connections appear to be the same as those that hold for other beliefs. What makes self-deception an interesting phenomenon for moral psychology is that much action that is morally wrong occurs out of self-deceptive *belief* that it is right. Thus, in adopting the belief view we accept a trade-off: we accept a psychological theory that makes better sense of the phenomenon of self-deception, but we are then burdened with the moral psychological project of explaining how and to what extent the self-deceived agent is responsible for the actions that spring from her self-deceptive belief.

Acknowledgments

This paper has benefited from helpful exchanges with Dagfinn Føllesdal, Wes Holliday, Krista Lawlor, Al Mele, John Perry, Tiffanie Poon, Ken Taylor, Robert Trivers, and two anonymous referees of *Erkenntnis*. I gave an early version to members of the Philosophy Department at Stanford in the spring of 2005. I completed this paper while supported by a fellowship from the Mellon Foundation.

Endnotes

¹ As Johnston (1988) notes, although Pears (1991) contests this objection.

² The word “sincerely” is tricky in this context. I’ll come back to it later.

³ This strategy may also seem to suffer from the problem that the *hide one of the beliefs somewhere else* strategy suffers from; namely, it’s not clear that anything like *deception* is happening.

⁴ Talbott (1995) has a stronger view than Mele (2001) on this matter. He holds that, since self-deception results in a belief, there *can’t* also be a belief in the contrary, i.e., doxastic alternative. Mele is open as to whether it’s possible for the self-deceiver to believe the doxastic alternative, but he holds that this isn’t the case in “garden variety” self-deception. (Note that “doxastic alternative” is my term; I’m simply using it here to help classify views.)

⁵ Funkhouser’s (2005) view is actually more complex than being a simple avowal view. I deal with it in this endnote so as not to cause digression from the main argument of this paper. Funkhouser claims that self-deception gives rise to avowals, which may seem to put him in the camp of the avowal theorists Audi and Rey. “Audi is correct that self-deceivers tend to say one thing, while truly believing otherwise. . . . the higher-order account explains why self-deceivers avow what they do, but nevertheless are guided in their actions by a contradictory belief” (pp. 307-308). But it is important to see that what Funkhouser views as self-deception actually *mixes* avowal and belief views on the product side. The characteristic form of self-deception for Funkhouser is that an agent desires to believe something (to make my point clear, I’ll abbreviate this: **D;B;p**) and then comes to believe *that she believes it* (**B;B;p**—hence the name “higher order account”); but then believing that she believes that *p*, she avows *p* without actually believing that *p*. Closer examination of this form will reveal that it is really just a sub-type of what I have called wishful self-deception, *since the content of the deceptive element and the content of the product are the same* (i.e., **B;p**—see section 2.1). So Funkhouser clearly holds that the product is a belief—just one with second-order contents—and the avowal may be seen as additional. Thus the disagreement between Funkhouser and myself is this: I think that the self-deception he identifies is one of a number of possible types, whereas he thinks it is the only type. I think examination of the range of cases discussed here supports my view.

⁶ This could be a desire, intention, or motivating emotion; of course, it’s controversial which of these it is and what its content has to be.

⁷ The spectrum on this front ranges from Mele, who holds that the agent merely has to have data available that supports the doxastic alternative, to those who hold that the agent simply *believes* the doxastic alternative. For a very instructive discussion of the possibilities here, see Pears (1984, p. 29).

⁸ You might claim that a disposition to avow is not a cognitive attitude, since dispositions can’t be regarded as true or false. But clearly, such dispositions involve some representation of things as being a certain way—they are not merely behavioral reflexes—and it’s fair to say that the content of such representations can be true or false. And the assertions they cause certainly will be.

⁹ Thus I will take it for granted that self-deception exists.

¹⁰ There are other logically possible options. But the avowal and belief possibilities are the dominant ones.

¹¹ It may be worth saying what notion of *content* I’m working with here. When I use “content,” I have in mind the classificatory notion of content that Perry (1998) explicates. That is, when I say that a mental state has a certain proposition as content, I mean that a proposition can be used to classify that state with others that are alike in significance.

¹² It is quite possible for a case of self-deception to be both wishful and willful, with both the desire that *p* and the desire to believe that *p* playing a causal role in subverting the agent’s epistemic norms to make self-deception come about. Nelkin (2002) holds the view that it is always a desire to believe that is operative in self-deception; I, of course, hold that this is just one type.

¹³ See Trivers and Newton (1982) for a discussion of this case.

¹⁴ Often there will also be a desire that matches the content of the product; that is, a case of self-deception can be *both* goal-driven and wishful (and possibly also willful). But it’s important to recognize the possibility of distinct types.

¹⁵ There is another prevalent type of mental phenomenon commonly called self-deception that I think would better be classified otherwise. The phenomenon I have in mind should be called *self-inflation bias*. This is simply a *general* tendency that many or most people have to see themselves in a comparatively more positive light than evidence would justify. This differs from the types of self-deception I discuss here

in that it is the result of a more general habit of thinking positively about oneself, as opposed to being caused by a specific desire or deceptive element. Thus, when people think about themselves as better drivers than they actually are, this is probably the result of self-inflation bias. The self-inflation bias seems to me to involve a fixed *frame* for interpreting incoming information in a self-flattering way. Self-deception, on the other hand, involves motivated selection of different frames. Furthermore, self-deception and self-inflation bias differ in that the former phenomenon, but not the latter, involves an epistemic tension in the mind of the agent between the evidence cognized and the product of self-deception; in the self-inflation bias, on the other hand, the agent is rather wholehearted in holding the flattering picture of herself. Unfortunately, no one I know of has distinguished clearly between self-deception and the self-inflation bias (cf. Taylor and Brown (1988)). This is not to say that the two phenomena are mutually exclusive. On the contrary, they can both be involved in the etiology of the same unjustified belief.

¹⁶ The idea that modulation of attention is involved in self-deception is developed in Talbot (1995), Mele (2001). I expand on the sketch I give here of how self-deception comes about in Van Leeuwen (forthcoming[a]) and Van Leeuwen (forthcoming[b]).

¹⁷ Rey doesn't talk at all about consciousness in his article.

¹⁸ See, for example, Funkhouser (2005) for such reasoning. This kind of reasoning doesn't actually seem to play a central role in Audi's or Rey's work.

¹⁹ There is much more to be said here about *how* the evidence that supports the doxastic alternative is encoded in the mind of the self-deceiver. This issue can't be decided here, but the following are some suggestions. Suppose an agent is self-deceived in believing that p . If this is the case, the agent doesn't believe that $\sim p$, but she does have beliefs in propositions (say, q , r , and s) that provide compelling evidence in favor of $\sim p$. We might call the beliefs that q , r , and s *caged* beliefs, for their subject matters are motivatedly not attended to; furthermore, *if* they were attended to adequately, they would produce the belief that $\sim p$. The self-deception is *precisely what prevents her from forming the belief that $\sim p$* and yields the belief that p . Nevertheless, a variety of sub-doxastic states (e.g., differing degrees of suspicion) can still result from the caged beliefs (i.e., her beliefs that q , that r , and that s) in the evidence that favors $\sim p$; it is the presence of these sub-doxastic states that gives rise to the epistemic tension inherent in self-deception. But, I must emphasize, adequately addressing this issue would require an entire paper on its own.

²⁰ See Lazar (1999) for a discussion of the role of emotions in self-deception.

²¹ That is, she finds the thought pleasant to run through her conscious mind, regardless of its truth. Such a process can lead to self-deception, but it is not tantamount to self-deception. An example of this kind of thinking would be the injured athlete who repeats in her mind the thought that her leg doesn't hurt; she knows it does, but this is a pleasant thought to have. Another possibility is that the woman in Rey's case does *believe* something self-deceptively, but what she believes is *that she believes* her lover is intelligent. Her self-deceptive belief is second-order, so we shouldn't expect it to have the consequences of a first-order belief simply that her lover is intelligent. This would make Rey's case conform to the form of self-deception that Funkhouser (2005) discusses. If this is the case, I think it would make more sense to say that she's self-deceived in believing *that she believes* he's intelligent. Thanks to one anonymous reviewer for suggesting this second possibility.

²² It should be clear that I take the pilot and the business woman to be *paradigm* cases of self-deception. Indeed, they do have the basic elements (intuitively appraised)—cognitive attitudes with content contrary to the weight of the evidence they've encountered, where those attitudes are held under the influence of motivations with related content. But, as one anonymous referee has pointed out, it is of course open to Audi or Rey to respond that these are not actually cases of self-deception. The reasoning would be that, since they have beliefs instead of avowals, they're not really self-deceivers. That would, I think, be an exceedingly puzzling move, since the fact that they actually believe makes it seem as if they are more seriously *deceived* than if they merely avow. And the source of the false cognitive attitude is the *self* in these cases just as much as it would be in cases that fit Audi's and Rey's definitions.

²³ One anonymous referee has pointed out that one must do more than believe evidence that contradicts an hypothesis in order to reject the hypothesis; one must also accept it *as* relevant contrary evidence. (E.g., believing that p won't get you to reject the hypothesis that q unless you also conceive of the truth of p as being contrary to the mooted truth of q .) This, however, doesn't undermine my point that beliefs govern hypotheses, since it will be *beliefs* about whether something is relevant evidence that matter. This does

show, however, that fully characterizing the governance relation will involve discussion of many more details than can fit into this paper.

²⁴ Velleman (2000) claims that there is no difference between the motivational role of beliefs and the motivational roles of other cognitive attitudes. I think this view, however, is based on wrongly generalizing from cases where believing and other cognitive attitudes cause similar behaviors to the conclusion that they have the same motivational role in all cases. My discussion in this section should make clear some of the ways in which motivational roles for cognitive attitudes differ.

²⁵ One way of distinguishing one cognitive attitude from another is by noting differences in kind of context in which they are psychomechanically effective. I pursue this strategy in forthcoming work.

²⁶ I have not discussed the relation between the product of self-deception and imagining here because the contents of imagining are too susceptible to other influences aside from beliefs, such as fears, so examining the contents of imagining is not likely to be illuminating in this context.

References

- Audi, R. (1982) "Self-Deception, Action and Will," *Erkenntnis* **18**(2): 133-158.
- Audi, R. (1988) "Self-Deception, Rationalization, and Reasons for Acting," in *Perspectives on Self-Deception*, edited by A. O. Rorty and B. P. McLaughlin, Berkeley, University of California Press: 92-120.
- Audi, R. (1997) "Comments," see Mele (1997).
- Borge, S. (2003) "The Myth of Self-Deception," *The Southern journal of philosophy* **41**(1): 1-28.
- Bratman, M. E. (1992) "Practical Reasoning and Acceptance in a Context," *Mind* **101**: 1-15.
- Davidson, D. (1982) "Paradoxes of irrationality," in *Philosophical Essays on Freud*, edited by R. Wollheim and J. Hopkins, Cambridge, Cambridge University Press.
- Davidson, D. (1985) "Deception and Division," in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, edited by E. LePore and B. P. McLaughlin, Oxford, Blackwell.
- Davidson, D. (1998) "Who is fooled?" in *Self-deception and Paradoxes of Rationality*, edited by J.-P. Dupuy, Stanford, CSLI: 1-18.
- Funkhouser, E. (2005) "Do the Self-Deceived Get What They Want?" *Pacific Philosophical Quarterly* **86**(3): 295-312.
- Johnston, M. (1988) "Self-Deception and the Nature of Mind," in *Perspectives on Self-Deception*, edited by A. O. Rorty and B. P. McLaughlin, Berkeley, University of California Press: 63-91.
- Lazar, A. (1999) "Deceiving Oneself or Self-Deceived," *Mind* **108**: 263 - 290.
- McLaughlin, B. P. (1988) "Exploring the Possibility of Self-Deception in Belief," in *Perspectives on Self-Deception*, edited by A. O. Rorty and B. P. McLaughlin, Berkeley, University of California Press.
- Mele, A. R. (1997) "Real self-deception," *Behavioral and Brain Sciences* **20**: 91-136 (with open peer commentary).
- Mele, A. R. (2001) *Self-Deception Unmasked*, Princeton, Princeton University Press.
- Nelkin, D. K. (2002) "Self-Deception, Motivation, and the Desire to Believe," *Pacific Philosophical Quarterly* **83**: 384-406.
- Pears, D. (1984) *Motivated Irrationality*, Oxford, Oxford University Press.
- Pears, D. (1991) "Self-Deceptive Belief-Formation," *Synthese* **89**: 393-405.
- Perry, J. (1998) *Reference and Reflexivity*, Stanford, CSLI Publications.
- Rey, G. (1988) "Toward a Computational Account of *Akrasia* and Self-Deception," in *Perspectives on Self-Deception*, edited by A. O. Rorty and B. P. McLaughlin, Berkeley, University of California Press: 264-296.
- Rey, G. (forthcoming) "Meta-atheism: Religious Avowal as Self-Deception," in *Philosophers Without God*, edited by L. Anthony, Oxford, Oxford University Press.
- Talbott, W. (1995) "Intentional Self-Deception in a Single Coherent Self," *Philosophy and Phenomenological Research* **55**(1): 27-74.
- Taylor, S. E., and J. D. Brown (1988) "Illusion and Well-Being: A Social Psychological Perspective on Mental Health," *Psychological Bulletin* **103**(2): 193-210.
- Trivers, R., and H. P. Newton (1982) "The crash of flight 90: doomed by self-deception?" *Science Digest* (November): 66-67, 111.
- Van Leeuwen, D. S. N. (forthcoming[a]) "Finite Rational Self-Deceivers," *Philosophical Studies*.

Van Leeuwen, D. S. N. (forthcoming[b]) "The Spandrels of Self-Deception: Prospects for a Biological Theory of a Mental Phenomenon," *Philosophical Psychology*.

Velleman, J. D. (2000) "On the aim of belief," in *The Possibility of Practical Reason*, edited Oxford, Oxford University Press: 244-281.