Self-Deception and the Limits of Folk Psychology

Eric Funkhouser

Abstract
This article considers the product of self-deception. Many assume, or argue, that the product of self-deception is a belief. I argue against this being a general truth by outlining some of the ways in which the self-deceived can be deeply conflicted, such that there is no fact of the matter concerning what they believe. These situations are not adequately addressed by many accounts of self-deception. Further, I argue that this task requires going beyond our folk psychological classifications.

1 Deeply Conflicted Self-Deceivers and a Static Problem

It would be nice if theorists interested in self-deception could start with an agreed upon understanding of the phenomenon — what it even means to be self-deceived and what are some of its paradigm examples — and then proceed from that starting point to address the various theoretical problems that it appears to pose. But, it should come as no surprise to any philosopher with any experience that there is no such shared understanding of what it means to be self-deceived. After examining many of the theories of self-deception and resolutions to particular problems that have been offered, it is clear that we are not all offering accounts of the same phenomenon. This point holds for many other philosophically interesting concepts as well, such as knowledge, free will, belief, rationality, etc. What we take to be problems and solutions in these areas, as well as convincing theories, depends on our initial understanding of the phenomenon. Problems, solutions, and theories are sometimes dismissed simply because they are judged to miss the target phenomenon. As we can imagine one saying: “What you say is fine as far
as it goes, I suppose. But that’s simply not what is meant by free will (or knowledge, or self-deception, etc.). You’ve given us a theory of something else.  

This much, however, is widely accepted: self-deception is some kind of motivated irrationality, in which the self-deceiver fails to handle the evidence available to her appropriately. Controversy arises when we try to specify the nature of this motivation and the resulting doxastic state of successful self-deception. These are two of the static problems of self-deception (in contrast to the dynamic problems which concern how this goal is achieved). In this article I will give particular attention to the doxastic question and consider, as Neil van Leeuwen has put it, what is the product of self-deception.  

We can start this project by considering what have been offered as paradigm examples of the phenomenon. Al Mele has offered the following:

For example, Sid is very fond of Roz, a college classmate with whom he often studies. Wanting it to be true that Roz loves him, he may interpret her refusing to date him and her reminding him that she has a steady boyfriend as an effort on her part to “play hard to get” in order to encourage Sid to continue to pursue her and prove that his love for her approximates hers for him. As Sid interprets Roz’s behavior, not only does it fail to count against the hypothesis that she loves him, it is evidence for the truth of that hypothesis.  

Mele goes on to describe Sid as believing that Roz loves him. In a previous article, I offered the following as a paradigm example of self-deception:

Nicole possesses much evidence that her husband Tony is having an affair with her friend Rachel. Nicole’s other friends have reported to her that Tony’s car is often seen parked in Rachel’s driveway, at times when he claims to be with his male friends. Tony has lost sexual interest in Nicole, and other suspicious behavior provides sufficient evidence for Nicole to be more than skeptical. Yet she laughs off the concerns of her girlfriends, and thinks to herself that Tony is certainly a faithful husband. (“After all, I am still an intelligent, charming, and attractive woman...”

\[\text{van Leeuwen (2007).}\]
\[\text{Mele (2001), p. 26.}\]
\[\text{Ibid., p. 46.}\]
— certainly more so than Rachel!”) Yet, in the evenings when Tony claims to be with his male friends, Nicole avoids driving by Rachel’s house — even when it requires her to drive out of her way.4

In that article I described this as a case in which Nicole does not believe that which she is motivated to believe (e.g., the proposition ‘Tony is not having an affair’), yet claimed that she was self-deceived anyway. I contrasted this example with what I called self-delusion, in which the subject clearly believes that which she is motivated to believe.

These examples, along with their descriptions, suggest conflicting answers to the static problem about the product of self-deception. In Mele’s example the self-deceived believe what (it seems) they want to believe, but according to mine they do not believe what they want to believe.5 The following possibilities then emerge:

1. At least one of these is not an example of self-deception.

2. Self-deception does not have a common product.

3. Self-deception does have a common product, it is just more sophisticated than we previously thought.

In that previous article I made the bold move of arguing for a version of 1, claiming that Mele offers an account of self-delusion but not self-deception. I still think that the distinction between self-delusion and self-deception, as made in that article, is important to theorizing about these cases, but I now think it was unwise and abusive to legislate usage such that Mele’s example does not count as self-deception. I wish to rectify that mistake, salvage what I feel was correct and insightful about that article, and consider the product question anew. My new answer falls under possibility 2, at least so long as we stick to traditional folk psychological concepts in our attempts to answer the product question.


5I should note, however, that Mele does not think that the self-deceived always believe what they want to believe. Further, he does not think that the self-deceived are always motivated by a desire to believe some proposition. Instead, Mele (2001) distinguishes between what he calls “straight” and “twisted” self-deception. As I understand his view, Mele holds that in many cases of twisted self-deception the self-deceived do not end up believing what they want to believe.
My main criticism of Mele’s account in Funkhouser (2005) was simply that there are cases of self-deception that fail to fit the account he offers. Further, I asserted that these cases presented the most philosophically interesting (because problematic) cases of self-deception all along. A bit rashly, I stipulated that the term ‘self-deception’ be reserved for these cases that Mele’s account missed — cases in which the agent is motivationally biased toward acquiring a certain belief, but she has only mixed success at fulfilling this goal. In cases like my above example there is a tension between some of Nicole’s behavior and her avowals and thoughts about her husband. Nicole does not behave as if she fully believes that her husband is faithful, though she claims both to others and herself that he is faithful. I now wish to introduce the technical term deeply conflicted self-deception to cover cases like these in which there is such non-trivial tension, such as between the self-deceived’s non-linguistic behavior and her avowals and private thoughts, so that it is unclear what the person really believes. Conflicts between avowals and non-linguistic behavior are stereotypical of self-deceivers, though this is just one form that the tension can take. Let self-deception be used to cover the broader category of motivated irrationality that covers both Mele’s paradigm cases (in which the false belief is clearly obtained) and my cases of deeply conflicted self-deception (in which it is unclear, if not indeterminate, what the person believes). In both Mele (2001) and in his more recent writings on self-deception (e.g., Mele (forthcoming)), Mele inevitably characterizes the phenomenon of self-deception as one in which the person is self-deceived in acquiring or retaining a belief. My main criticism of Mele’s treatments of self-deception, then, continues to be that they do not address the doxastic condition of deeply conflicted self-deceivers. And many instances of self-deception are of this type. Whether or not this is a significant objection to Mele, as it may simply be a phenomenon that he was not targeting, I will argue that such cases are naturally understood as examples of self-deception, have great theoretical interest and shed light on the limits of belief.

6See Audi (1985, 1988) and Funkhouser (2005) for accounts of self-deception that see the avowal/behavioral tension as characteristic of the phenomenon.
2 The First Step Toward a Solution: Deconstructing Belief

The product question for deeply conflicted self-deception is as follows: What do such people really believe (during their time of deep conflict)? This question raises a problem because for the deeply conflicted different belief-indicators, such as their behavior and their avowals, oppose one another (i.e., one belief-indicator counts in favor of attributing a belief that \( p \) and the other counts against this attribution over a proximate, if not simultaneous, time span). Further, in at least some cases such opposing indicators are both prominent and persistent in the person’s decision making. For example, both Nicole’s avoidance behavior while driving as well as her public and private avowals could be a routine and steady part of her life. So, what does Nicole really believe concerning the fidelity of her husband?

The problem with this question, I suggest, is that it falsely presupposes that there is a fact of the matter as to what Nicole really believes. But, the question is useful anyway. For, it exposes the limits of the applicability of such folk psychological classifications. Further, considering the limits of belief can prompt us to make the necessary revisions for a more accurate psychology.

I think we are doing justice to common usage, as well as to philosophical and psychological theorizing, if we characterize belief as the psychological attitude in which a person regards a proposition as true.\(^7\) This characterization still leaves significant room for debate. Namely, what is it for a person to regard a proposition as true? This is where disagreement emerges. Some will say that a person regards a proposition, \( p \), as true if she acts (or is disposed to act) to satisfy her desires as if the world were such that \( p \). Others will say that a person regards \( p \) as true if she stores a token of \( p \) in her “belief-box”.\(^8\) And so on. While different theorists will offer competing accounts of what it is to regard a proposition as true, I am optimistic that, if queried, there would be wide agreement among them that a person believes a proposition simply if she regards that proposition as true.

\(^7\)Others, such as Velleman (2000), would add that belief is the psychological attitude of regarding a proposition as true with the aim of getting the truth value right. I argued against this further qualification in Funkhouser (2003).

\(^8\)Braithwaite (1932/1933) is representative of the former, dispositionalist, option. And Fodor (1975) offers the classic statement of the “belief-box” option.
In making these claims about belief I have been careful to use the word ‘person’. This is because beliefs are attributable to the person as a whole, and as such they are *personal* attitudes. Belief is also the most general regarding-as-true stance — to believe a proposition is to regard it as true all things (i.e., psychologically relevant factors) considered. This characterization of belief suggests the possibility of interesting contrasts: *sub-personal* regarding-as-true stances and *qualified* regarding-as-true stances. Sub-personal regarding-as-true stances are those that are attributable, in the first instance at least, only to subsystems of a person. For example, a person acquainted with the Muller-Lyer illusion might believe that the two lines before him are of equal length, though his visual system “believes” otherwise. It is natural to use the expression “believes” here, I think, because the visual system can be seen as regarding a proposition as true (i.e., the proposition ‘One of these lines is longer than the other’). A qualified regarding-as-true stance, in contrast, may be appropriately attributed to the person as a whole, but only with respect to a certain aspect. For example, we may say that, with respect to her driving behavior, Nicole regards it as true that her husband is unfaithful. It would be a mistake, though, to describe Nicole’s behavior as one of her subsystems. Rather, her behavior is one aspect by which she can, among other things, regard a proposition as true. Nicole’s driving behavior, we can say, stands in the regarding-as-true relation or stance toward the proposition ‘My husband is having an affair with Rachel’ (given, of course, Nicole’s background desires and beliefs).

Let’s now consider, with greater detail, the various categories of regarding-as-true stances that count as belief-indicators. The list below might not be exhaustive, but it is certainly a good start. Further, I will provide an example of each category of regarding-as-true stance that is relevant to my original example of self-deception. These examples will help shed light on the various ways in which the self-deceived, or anyone else, can be deeply conflicted.

1. **Theoretical Reasoning.** When we engage in theoretical reasoning we regard some propositions as true and others as false. This is to accept or reject propositions as premises in arguments where the outputs are other propositions that are regarded as true. Let us suppose that in her theoretical reasoning Nicole accepts the premise that all adulterers are morally reprehensible. However, when considering the morality of her husband in relation to adultery, she does not reach the judgment that Tony is morally reprehensible. This is not because Nicole is horrible at simple logic. Rather,
she simply does not regard the proposition ‘Tony is an adulterer’ (or ‘Tony is having an affair with Rachel’, etc.) as true for the purposes of this argument.

It should be pointed out that a proposition can be regarded as true for the purposes of theoretical reasoning in some contexts but not in others. For example, it is possible for one to regard a proposition as true when used as a premise for one class of arguments, but not regard it as true when used as a premise for other arguments. (In extreme cases, we regard a proposition as true for the purposes of only one theoretical argument. E.g., in Euclid’s proof for the infinity of primes we begin with a premise that assumes there is a largest prime.) The purpose of the reasoning, the nature of the subject matter, its importance to the agent, and the environment in which one finds oneself when reasoning about that subject matter can all contribute to this effect. Here think of the different propositions that some regard as true when reasoning about religious matters or their friends, but they do not regard as true when reasoning about other matters. This context-relativity and compartmentalization also holds for the other regarding-as-true stances discussed below.

2. Practical Reasoning. We also regard propositions as true or false for the purposes of practical reasoning, where the output is an action or intention to act. Significantly, the propositions that we regard as true for practical reasoning need not match those that we regard as true for theoretical reasoning. In our example, Nicole avoids driving her car by Rachel’s house, guided by practical reasoning that regards as true the proposition ‘Tony is having an affair with Rachel’, while nevertheless concurrently rejecting that proposition for the purposes of theoretical reasoning about Tony’s moral standing (as discussed above).

3. Behavior. Even when we are not engaged in any practical reasoning, our behavior by itself can constitute a regarding-as-true stance toward a proposition. This behavior can be subtle, such as the body language Nicole uses with Tony and Rachel. It can also be ambiguous — e.g., Do her excessive displays of physical affection in public indicate that she regards it as true that Tony has been faithful (and, hence, worthy of her affection), or do they indicate the opposite (employed, say, as an attempt to seduce him back)? But this behavior can also be neither subtle nor ambiguous — such as when Nicole shoots a heated glare at Rachel. The behavior in this particular example is not the product of practical reasoning, but goes at least some way toward indicating that she regards it as true that Rachel has been with Tony.

4. Internal Reports. We can also regard a proposition as true simply in
the sense of reporting it as true to ourselves — e.g., by an act of introspection or internal assent. Of course, such reports carry more weight as a belief-indicator when they are sincere and reflective. Nicole internally reports to herself (perhaps sincerely and reflectively) that Tony is faithful, and thus regards that proposition as true in at least this limited sense.

5. External Reports. An external report is a public affirmation of a proposition, typically done through language (though, perhaps there are other means of communicating these affirmations). Nicole reports (again, perhaps even sincerely and reflectively) to her girlfriends that Tony is faithful. Here the external report matches the internal report, but this is not necessary even when each is sincere and reflective.

6. Emotion. Some emotional responses, especially those to which some norm of rationality is applicable, also count as regarding-as-true stances toward a proposition. For example, fearing a snake is one (limited) way in which one can regard it as true that the snake is dangerous. As with all other regarding-as-true stances, these emotional responses can be outweighed by other regarding-as-true stances such that, all things considered, perhaps our fearful subject does not regard it as true that the snake is dangerous.

Undoubtedly many self-deceivers have emotional responses that contribute to these “all things considered” judgments. Nicole might experience negative emotional reactions toward both Rachel and Tony — e.g., resentment and jealousy — that, in themselves, count as regarding-as-true stances toward the proposition ‘Tony has been unfaithful (with Rachel)’. For example, in the act of resenting Rachel, Nicole manifests regard for the truth of their infidelity. This is all compatible with Nicole experiencing other emotions which point in the opposite direction. Let us suppose that Nicole also feels steadfast devotion toward Tony, even though she has a long-standing intolerance of adultery. Even if this particular example is not convincing, it should be clear that there can be conflicting regarding-as-true stances even within a common category (such as emotion).

7. Perception. Finally, we can regard a proposition as true simply in virtue of having a perceptual experience. Here, the claim is, the perceptual experience itself counts as a regarding-as-true stance independent of any contribution this perceptual experience goes on to make toward our reasoning, reports, etc. Again, even one who is aware of the Muller-Lyer illusion can perceptually regard it as true that one line is longer than the other. Or, in an extreme situation, Nicole could even perceive Tony’s infidelity with Rachel (i.e., catch them in the act) and yet fail to believe (i.e., regard-as-true, all
things considered) that Tony has been unfaithful. Seeing isn’t always believing.

I submit that there are at least these 7 different categories of regarding-as-true stances. These different categories of stances can oppose one another, as the above examples illustrate. Further, there can even be opposition within a category both across contexts (e.g., a proposition is regarded as true for the purposes of theoretical reasoning about religious, but not other, matters) and within a given context (e.g., a proposition is regarded as true by one sense modality, but not another, in the same environment). This allows for many permutations of deep doxastic conflict in general, and self-deception in particular.

There is, I confess, an additional assumption at work in reaching this conclusion that there are many varieties of deeply conflicted self-deceivers. Namely, I am assuming that a person’s beliefs are constituted by their status with respect these different regarding-as-true stances, rather than these being mere indicators or evidence of their belief state. In this sense, I claim that belief reduces to, or is nothing over and above, these regarding-as-true stances. I will not argue for this assumption here, but I think that a bit of reflection will show that it has high plausibility. For example, it is difficult to see how it is conceptually possible for one to regard something as true for the purposes of theoretical reasoning, practical reasoning, etc. but yet not believe it.

By taxonomizing the various regarding-as-true stances we can see the many possible forms that deeply conflicted self-deception can take. It has been common to focus on a few of these — e.g., the conflicts between avowals and certain behaviors. But the possibilities are much richer than that, and there is no reason to think that self-deception results in only a narrow band of regarding-as-true conflicts. We should expect there to be cases of self-deception that manifest deep conflicts between, say, emotional responses and behavior, theoretical reasoning and perception, etc. Of course, these conflicts are typically much more complex, as they inevitably involve more than just two categories of regarding-as-true stances. But some of these conflicts between paired stances simply are more salient.
3 The Second Step Toward a Solution: No Privileged Weighting

Remember the contrasting cases of self-deception offered by Mele and myself, discussed above. In Mele’s example Sid acquires the belief that he is motivated to acquire (at least, Mele describes Sid as acquiring this belief), whereas in my case Nicole’s belief state is less clear and she appears to have mixed success in acquiring this belief. I want to claim that this lack of clarity is not due to our limited epistemic perspective, rather it is a real indeterminacy in the world. There is no fact of the matter as to what Nicole believes.

Belief is a rough and ready folk-psychological concept that is very helpful in predicting and explaining behavior in ordinary circumstances. But given that it is a folk concept, there is no reason to think that there are strict rules for its application such that there always is a determinate fact of the matter whether a given person believes that \( p \). How could things be otherwise? If we accept that what a person believes is constituted by her status with respect the regarding-as-true stances discussed above and we also allow for significant conflict among such stances, then there is a fact of the matter as to what each person believes, with respect to each proposition, only if there is a fact of the matter as to how much weight should be given to each regarding-as-true stance.

My previous position (Funkhouser, 2005) on cases of deeply conflicted self-deceivers, such as Nicole, was that their behavior concerning important matters, in particular behavior aimed at avoiding situations that would jeopardize their goal of avoiding the truth (e.g., situations, perhaps, such as seeing the car in the driveway), points to them believing the truth “deep down” all along. Part of this position was the view that behavior, either all things considered or in certain high stakes situations (actual or hypothetical), was the decisive indicator of real belief. I no longer accept this position, as I think that I underestimated, in a couple regards, the theoretical difficulties

---

9Schwitzgebel (2001) provides some additional reasons, besides my “no privileged weighting” claim of this section, for which a person could be in a state of “in-between” believing.

10For example, I wrote: “Again, the view of belief favored here is one which places greater emphasis on non-linguistic behavioral dispositions than first-person phenomenology and linguistic reports. And the behavioral dispositions of the self-deceived, especially in situations where the costs of mistake are high, are tipped toward believing the truth” (Funkhouser, 2005, p. 307).
posed by deeply conflicted self-deception.

First, I falsely assumed that the conflict would also be between behavior and avowal. But, this is not always the case. We have already imagined situations, which I assume are also psychologically real, in which the conflict is between, say, emotional responses and theoretical reasoning (here, assume that the behavior is neutral). In order for there to be a determinate fact of the matter concerning belief there would have to be a way to settle all such conflicts among regarding-as-true stances. Second, I no longer accept my earlier claim that non-linguistic behavior, particularly in high stakes contexts, trumps avowals and determines what a person really believes. I was imagining that if push came to shove — that is, if the costs of neglecting the truth became too great — the deeply conflicted would behave so as to respect the truth and that this would show that they believed the truth all along. But, there are a couple problems with this answer. One problem is that once the person finds herself in this high stakes environment it may be the case that she suddenly realizes the truth for the first time, rather than manifesting a psychological attitude that she had all along. Further, even if her behavior comes to present a unified regarding-as-true stance it need not be the case that her theoretical reasoning, practical reasoning, etc. will then fall in line as well. So, there could still be regarding-as-true conflicts between behavior and, say, theoretical reasoning and internal reports. And, if belief simply is the most general regarding-as-true stance (i.e., regarding-as-true, all things considered), then I see no reason to think that a person who behaves as if \( p \) (given her desires and background beliefs) thereby believes that \( p \), when she fails to engage in theoretical reasoning as if \( p \) (and might even theoretically reason as if not-\( p \)) and fails to internally report \( p \) (and might even internally report not-\( p \)).

Recall that we were led to this conclusion of real indeterminacy for the deeply conflicted based on two assumptions: 1) Belief is constituted by the values for various regarding-as-true stances taxonomized above, and 2) There is no privileged weighting of these values that determines “real belief”. Nevertheless, in many pedestrian cases it is clear (determinate) what a person believes. But this is because all the plausible candidates for weighing the different regarding-as-true stances converge on the same result. We should not be led, by the phenomenon of indeterminacy based on conflict, to an eliminativism about belief. Instead, we should only recognize its limits of application. Of course, it very well might be the case that, so long as these sub-personal and qualified regarding-as-true stances are determinate, there
is a determinate fact of the matter as to whether a person believes that \( p \) according to a given candidate for weighing the various regarding-as-true stances. But, the central claim of this section is that there is no good reason to think that one of these candidates best fits our folk concept of belief. In fact, persistent disagreement over how to describe many psychological subjects, and self-deceivers in particular, supports the claim that no single weighting is privileged by our ordinary practices. (I am also skeptical that any particular weighting is privileged by psychological/theoretical practice. Rather, I imagine that different weightings are preferred for different theoretical/explanatory tasks.)

4 The Third Step Toward a Solution: Psychological Descent

If these central claims about belief are correct, then the lesson to be learned, with respect to the product question for self-deception, is that we must descend to a lower-level (i.e., the component regarding-as-true stances) in order to adequately characterize the cognitive states of at least some of the self-deceived. The only other alternative is to rest content with an answer such as, “It is simply indeterminate what the person believes in these situations. The self-deceived are deeply confused or conflicted.” This answer is true, but we can be more informative.

It is noteworthy that the same psychological descent that I recommend for the product question might also be desirable when characterizing the motives of the self-deceived. I previously argued that the self-deceived are/were motivated by a desire to acquire a certain belief. I put the point as follows:

\[
\ldots \text{the self-deceiver is guided by a desire for a certain state of mind. Perhaps the desire can be given greater specificity in certain cases, but the most appropriate generalization we can truthfully make of self-deceivers, limiting ourselves to the traditional categories of folk psychology, is that they are guided by a desire to believe.}^{11}
\]

This still seems right, but with our deconstruction of belief at hand we can

---

\(^{11}\)Funkhouser (2005), p. 298. Also see Nelkin (2002) for argument for the desire-to-believe account of motivation in self-deception.
now recognize some of the more particular “states of mind” that the self-deceived can be motivated to acquire. Rather than desiring belief *simpliciter*, the self-deceived might be motivated to acquire specific values for one or more of the component regarding-as-true stances. In particular, it might be common for them to be motivated to make certain internal and external reports (and/or acquire the disposition to so report). This could explain why it is stereotypical for self-deceivers to avow one thing while exhibiting salient behavior that counters the avowal.

One might worry, however, that such nuanced desires are psychologically unrealistic. After all, belief is the folk concept, not these component regarding-as-true stances. And one might object to this more nuanced account of the self-deceived’s desires on the grounds that ordinary folk — the majority of the self-deceived — cannot have desires with this fine-grained content if they do not possess concepts for these component regarding-as-true stances. The mistake in this objection, however, is in the claim that the ordinary folk do not possess concepts corresponding to these various stances. Ordinary people might not think of belief as being constituted by these various stances, but they certainly have the concept of reporting something (to oneself or others) as true, having particular emotional responses (that, on reflection, they would recognize as appropriate only given that certain conditions hold), perceiving something as being the case, etc. And, they can certainly desire to be in particular psychological states that fall under these categories. It is not psychologically unrealistic to suppose that Nicole desires to report, to herself and others, that Tony is faithful or that she desires to feel steadfast devotion for him.

Here is a more straightforward way of arguing for the claim that the self-deceived have desires that are more sophisticated than our crude, folk-psychological approximation that they desire to believe that $p$: There are examples of self-deception in which the person is wholly successful with respect to her goal of deceiving herself, yet she does not believe that $p$. These are cases of deeply conflicted self-deception. But if this self-deception is wholly successful, then the desire that motivates it has been satisfied. Since the person does not believe that $p$, the desire that motivated this self-deception could not, then, be a desire to believe that $p$. (Instead, the desire is for some psychological state that the person actually enters into.)

I anticipate that two types of objections to this argument will be common. First, some will simply deny the claim that the deeply conflicted do not believe that $p$. I have already countered this claim in the previous section.
Second, others will accept my claim that the deeply conflicted do not believe that \( p \), but they will claim that, for this very reason, the self-deception is not fully successful. More radically, they may even say that these are not cases of self-deception at all. Yet, I find each of these claims implausible. Consider again our Nicole, who sincerely and reflectively avows both to others and herself that Tony is not having an affair. Yet, she carefully avoids driving by Rachel’s house at times during which an objective observer would realize that Nicole would have good reason to believe that Tony might be there. We could add to the avoidance behavior in both its range and detail (e.g., she goes out of her way to avoid looking in the pockets of his clothing if she happens to do the laundry) in order to enhance the conflict between her avowals and behavior. Let us suppose that this is a case of deep conflict and that we were correct in concluding that there is no fact of the matter as to what she believes concerning Tony’s fidelity. In this case, it is implausible to hold that her self-deception would be more successful were she to lack such avoidance behavior. Yet if this behavior were lacking the conflict would dissipate, and it would be clear that she believes that \( p \). But such a state would likely frustrate her goal. For Nicole, like most people, is not completely irrational and insensitive to evidence. If she were to confidently drive by Rachel’s house and search through her husband’s clothes with the confidence of one who fully believes in his fidelity, she would likely discover the disturbing truth. And then she could no longer avow to others, or at least to herself (which is probably more important), that Tony has been faithful. So in some sense — i.e., the behavioral (and probably practical reasoning) regarding-as-true sense — she knew all along that Tony has been unfaithful. Of course, I have argued that in another sense — i.e., the internal and external reports regarding-as-true senses — she “knew” otherwise. If it is true, as it now appears upon recognizing the frustration she would experience were she to behave a true believer, that her goal is to preserve these avowals, then it seems that her deception is wholly successful. She is not motivated to believe that \( p \) simpliciter, but only to regard \( p \) as true in the internal and external reporting senses. And, the case so described and analyzed, I certainly cannot see how one could think that this is not self-deception at all.

Let us now summarize the main conclusions regarding the product of self-deception. I have, retracting my previous stipulations, agreed to use the term ‘self-deception’ liberally so as to cover the types of cases presented by Mele (and others) in which it seems clear that the self-deceived acquire
the belief that $p$. I call these cases of self-delusion. But these are only a proper subset of the cases of self-deception. I have argued that there are other cases in which the self-deceived produce a confused belief-like condition, so that it is genuinely indeterminate what they believe with respect to $p$. Yet, such people, the deeply conflicted, can be wholly self-deceived nonetheless. If this is true, then self-deception does not produce a common product. As a corollary, given the claim that the deeply conflicted can be wholly successful *qua* self-deception, self-deception does not have a common motivation either. To characterize the motives and products of the deeply conflicted we must abandon our simple folk classifications and descend to the level of the regarding-as-true stances described in section 2. The protracted debates over the motives and products of self-deception are largely due to disputants using folk psychological tools that simply are not up to the task. Self-deception is interesting, among other reasons, because it helps reveal the limits of folk psychology and suggests more sophisticated ways to analyze our psychological attitudes toward propositions.

References


