

# **LIBERALIZING SELF-DECEPTION: REPLACING PARADIGMATIC-STATE ACCOUNTS OF SELF-DECEPTION WITH A DYNAMIC VIEW OF THE SELF-DECEPTIVE PROCESS**

Patrizia Pedrini

Volume 13, Number 2, Summer 2018

URI: <https://id.erudit.org/iderudit/1059497ar>

DOI : <https://doi.org/10.7202/1059497ar>

[See table of contents](#)

Publisher(s)

Centre de recherche en éthique de l'Université de Montréal

ISSN

1718-9977 (digital)

[Explore this journal](#)

Cite this article

Pedrini, P. (2018). LIBERALIZING SELF-DECEPTION: REPLACING PARADIGMATIC-STATE ACCOUNTS OF SELF-DECEPTION WITH A DYNAMIC VIEW OF THE SELF-DECEPTIVE PROCESS. *Les ateliers de l'éthique*, 13, (2), 11–24. <https://doi.org/10.7202/1059497ar>

Article abstract

In this paper, I argue that paradigmatic-state accounts of self-deception suffer from a problem of restrictedness that does not do justice to the complexities of the phenomenon. In particular, I argue that the very search for a paradigmatic state of self-deception greatly overlooks the dynamic dimension of the self-deceptive process, which allows the inclusion of more mental states than paradigmatic-state accounts consider. I will discuss the inadequacy of any such accounts, and I will argue that we should replace them with a dynamic view of self-deception that is more liberal regarding the mental states in which self-deceivers may find themselves.



# LIBERALIZING SELF-DECEPTION: REPLACING PARADIGMATIC-STATE ACCOUNTS OF SELF-DECEPTION WITH A DYNAMIC VIEW OF THE SELF-DECEPTIVE PROCESS

PATRIZIA PEDRINI  
UNIVERSITY OF FLORENCE, ITALY

## ABSTRACT:

In this paper, I argue that paradigmatic-state accounts of self-deception suffer from a problem of restrictedness that does not do justice to the complexities of the phenomenon. In particular, I argue that the very search for a paradigmatic state of self-deception greatly overlooks the dynamic dimension of the self-deceptive process, which allows the inclusion of more mental states than paradigmatic-state accounts consider. I will discuss the inadequacy of any such accounts, and I will argue that we should replace them with a dynamic view of self-deception that is more liberal regarding the mental states in which self-deceivers may find themselves.

## RÉSUMÉ :

Dans cet article, je soutiens que les explications de l'auto-illusion en termes d'un état « paradigmatique » souffrent d'un problème de limitation qui ne rend pas justice à la complexité du phénomène. Plus précisément, j'avance que la recherche même d'un état paradigmatique néglige tout à fait la dimension dynamique du processus d'auto-illusion, de sorte à inclure davantage d'états mentaux que ne le font les explications en termes d'un état paradigmatique. Après avoir démontré l'insuffisance de ces dernières, je proposerai que nous devrions les remplacer par une conception dynamique de l'auto-illusion qui serait plus flexible quant aux états mentaux potentiellement vécus par des personnes sous l'emprise de l'auto-illusion.

## 1. INTRODUCTION

The problem of the mental state in which self-deceivers find themselves has a long tradition. It can be traced back at least to Mele's early objections (1997, 2001) to Davidson's intentionalism (1985). Famously, Davidson's idea that self-deception must be intentional was criticized by Mele for leading to a couple of paradoxes, one of which is the "static paradox," as it is known.<sup>1</sup> It regards the mental state a self-deceiver is in when the self-deceptive process is successfully accomplished. Mele's argument is well known: if self-deception is intentional, and self-deceivers thus intend their self-deception, at the end of the self-deceptive process they must retain the belief that not- $p$ —that is, the belief about how things really stand—while also getting to believe the self-deceptive, desired falsity that  $p$ . If this is correct, then it seems that the final, resulting mental state in which self-deceivers find themselves is somehow paradoxical, amounting to both believing that  $p$  and also believing that not- $p$ .<sup>2</sup> Thus, Mele suggested that we should get rid of intentionalism altogether and resort to focusing on the motivational set of the subjects engaging in self-deception: since the subjects desire that not- $p$ , the desire in question biases their evaluation and selective search for evidence. This opens the door to a motivationally distorted treatment of the relevant data, which leads the subjects directly into believing that not- $p$ , without also retaining any belief that  $p$ .

However, although the solution offered by Mele may be a satisfactory way out of the static paradox, it has the drawback of describing the state of mind of the self-deceiver as quite peaceful: if a full-blown belief that  $p$  is successfully reached, then there is no trace of the psychological tension that seems, instead, to be highly typical of self-deception. For this tension is obviously due to the fact that the motivationally distorted self-deceptive process runs counter to evidence that is at hand or easily available that not- $p$ .

Of course, scholars are sensitive to the interplay of motivation and evidence, and to their opposing thrusts. And once the problem of the psychological tension created by this contrasting interaction has been posed as crucial, scholars have continued to investigate the nature of this resulting, final state of mind of self-deception, and have tried to keep psychological tension in the picture. To this end, they have advanced interesting and refined analyses of what this final state must be, virtually all requiring that a satisfactory account of self-deception must preserve and account for the psychological tension that is characteristic of the self-deceiver's final state of mind.

In this paper, I wish to raise the issue, again and afresh. I will argue that, while most accounts currently on offer are on the right track in their search for states of mind that can account for tension, I will also object that most of them are on a misleading track insofar as they look, or tend to look, for a paradigmatic, final mental state for self-deception. For virtually all accounts working with paradigmatic states for self-deception suffer from a potential flaw that should be carefully considered—namely, a certain restrictedness, at least in spirit. I say "at

least in spirit” because, if a state is presented as paradigmatic, that need not be necessary. Thus, there may be room for predicting other mental states in a descriptive account of the mental state a self-deceiver is in—a less paradigmatic one, say, and yet capable of satisfying the constraints set by the motivated self-deceiver’s struggle with opposing evidence. This seems to be especially true for all those accounts that offer only sufficient (and not also necessary) conditions for self-deception, such as Mele’s.<sup>3</sup> However, the rhetoric of most accounts addressing the problem of the paradigmatic mental state of self-deception may lead us to assume that, when self-deceiving, more often than not we are in a certain highly typical state, and to focus on that state of mind at the expense of other possibilities. My plan in this paper is to show that these other possibilities are not only live, but also quite common and highly typical as well. They are so common and typical that, as a matter of fact, once we have and keep this empirical truth in view, any claims regarding states that can be taken as paradigmatic become fatally weakened.

One might ask why there has been such a great focus on the problem of a paradigmatic state. Most likely, it is created by a bias in favour of a static state, which questionably guides the analysis toward what I will dub a “snapshot theory” of self-deception. Unpacking the metaphor, one discovers that a snapshot theory of self-deception seems to be designed to take a descriptive, static picture of the mental state a self-deceiver is in. But this focus on static states may be highly misleading and lead us to exclude other candidates for typical self-deceptive mental states.

Starting from these premises, I will argue that self-deception is a psychological process before it is an end state—a process set in motion by the force field created by motivation and evidence. Accordingly, I will offer an argument in favour of replacing what I call “paradigmatic-state accounts” of self-deception with a dynamic view of the self-deceptive process. Once we have seen the reasons in favour of this move, and once the move is made, we will be in a position to see that mental states that are taken to be paradigmatic of self-deception should be liberalized, so as to include all the variations allowed by the evolving combinations of the two factors of motivation and evidence. I will focus fully on motivation and evidence as the two fundamental constraints singling out the phenomenon of self-deception. I will explain how the pull of motivation and the thrust of reality create a force field that is dynamic, often in motion, strained by the variations in these two components that may be triggered by various contingent or noncontingent factors. If we bear it in mind that self-deception amounts to a dynamic psychological space determined by both motivation and evidence, then a dynamic view becomes a promising option. And this significantly changes our approach to the attitudes that are the products of the self-deceptive process. Such a move turns out to be liberalizing regarding the mental states that can be instantiated in self-deception and whose possibilities should be brought fully into the picture.

Here is the plan of the paper. In section 2 I will discuss the inadequacy of the paradigmatic-state accounts of self-deception, thus creating the premises for moving forward toward formulating an alternative view. In section 3 I will explore the most promising alternative—namely, a dynamic, more liberal theory of the self-deceptive process, which amounts to my proposed positive view. I will also address one obvious objection to my positive view, and I will conclude by briefly indicating what my approach suggests in terms of refining and applying the conceptual mental categories that prove to be useful for capturing self-deception.

## 2. THE INADEQUACY OF PARADIGMATIC-STATE ACCOUNTS OF SELF-DECEPTION

As we have seen, the problem of including and explaining psychological tension in a satisfactory account of self-deception has led several scholars to advance proposals suggesting that we should adopt an “attitude adjustment approach” (cf. Deweese-Boyd, 2017) regarding the mental state that is the product of self-deception. Since full-blown self-deceptive belief as a final product of self-deception seems unable to explain why the self-deceiver experiences psychological tension, we can choose to adopt alternatives. One possible alternative is to posit some quasi-doxastic or other nondoxastic attitudes towards the self-deceptive proposition. Candidates are hopes, suspicions, doubts, anxieties (Edwards, 2013), “besires” (Egan, 2009), pretense (Gendler, 2007), and imagination (Lazar, 1999). On this view, the subject is not required to end up believing a desired proposition. Rather, the subject can either entertain the hope that the desired proposition is true or suspect that the desired proposition is not true, or the subject can have doubts about its truth value or maybe also anxiously fear the possibility of it being false. All these attitudes seem to be able to explain why the self-deceiver is in a highly tensive state of mind: since the evidence points to a certain truth value of the desired proposition—that is, it suggests that it is at least likely that it is false—the subject struggles with such evidence and tries to see if the desired proposition can be true.

Another alternative is to alter the content of the proposition believed. We can do so without incurring any static paradox of the kind associated with the traditional intentionalist models of self-deception. For instance, Funkhouser (2005) suggests that self-deceivers have a second-order belief about their believing that  $p$ , while they do not believe that  $p$  at all. Reality shows them that  $p$  cannot be true or that  $p$  is unlikely; however, they process the evidence in a motivationally biased way, so that they can at least believe that they believe that  $p$ .<sup>4</sup> That creates a tension between the false second-order belief that they believe that  $p$ , along with the dispositions associated with having it, and the first-order belief that not- $p$ , along with the dispositions associated with it. Bilgrami (2006) also suggests that the tension is due to a conflict, this time between a fully authoritative, true, second-order belief about a completely transparent, first-order belief that  $p$ , and another first-order belief that not- $p$  that is, however, not transparent, and that does not generate any second-order belief that one also has the first-order belief

that not- $p$ . Here the tension is created either by the conflict between the transparent, first-order belief that not- $p$  and the opaque, first-order belief that  $p$ , or by the clash between the true, second-order belief that one has a first-order belief that  $p$  and the first-order belief that not- $p$ , or both, along with the complications created by the further dispositions associated with all of these beliefs.

More recently, Lynch (2012) has claimed that “wholeheartedly believing what one wants to be true may be rare in self-deception (p. 440), and that we should look for a “more fine-grained way of capturing the attitudes of subjects towards propositions than can be accomplished with the coarser apparatus of belief” (p. 438). Thus, he claims that unwarranted degrees of confidence in  $p$  are enough to explain tension.

He also argues that we have self-deception as long as the subject puts some different degree of confidence in  $p$  and in not- $p$ , whereas any phenomena in which the subject avoids altogether the question whether  $p$  are best captured as “escapism” (Lynch, 2012, p. 446). He draws on Longeway (1990) and describes escapism as a defence against reality. According to Lynch, “deep conflict cases” best represent escapism: while in self-deception there is a cognitive tension due to the different degrees of confidence placed by the subject in  $p$  and not- $p$ , in deep conflict cases we have a more profound tension that is behavioural. The subjects here are taking a greater risk by acting upon their attitudes than they take by just speaking and thinking (p. 435). Often the subjects engage in avoidance behavior. Such actions give us reasons to think that the subjects are not simply struggling cognitively with  $p$  and not- $p$ , but that they are investing in  $p$  in a way that shows that they must have found a way to avoid any vacillation as to whether  $p$ . According to Lynch, this may not be self-deception any longer—or at least not a paradigmatic case of it. Interestingly, however, he admits that it may not be a necessary truth that self-deception contains tension, and allows for the possibility that, at times, the subject may become fully convinced of the desired proposition that  $p$  (p. 442). I will come back to this later in the paper.

In the context of distinguishing willful ignorance from self-deception, Lynch (2016) is even more explicitly interested in singling out paradigmatic cases of self-deception. He is aware that “philosophical analysis of a phenomenon is challenging enough at the best of times, but it becomes all the more difficult when there is disagreement over what the paradigmatic cases are. Such is the situation, unfortunately, with regards to self-deception” (p. 513). However, he thinks that “there are some features that are generally recognized to be present in paradigmatic self-deception,” such as

- (1) the subject’s encounter with evidence indicating that some true proposition, not- $p$ , is true; and
- (2) the strong desire of the subject that  $p$  be true.

So, according to (1) and (2), in paradigmatic self-deception the subject encounters *unwelcome* evidence, indicating that not- $p$ . Lynch adds that “beyond that, disagreement persists, particularly with regard to the subject’s epistemic/doxastic relation with the truth” (p. 513).

Lynch then claims that approaches regarding paradigmatic cases of self-deception can be sorted into three categories (p. 513-514):

- (a) unwarranted-belief accounts, where the subject ends up self-deceptively believing that  $p$  (Mele, 1997, is an example of such an account);
- (b) implicit-knowledge accounts (e.g., Bach, 1981), where the subject does not believe that  $p$ , but recognizes the truth of not- $p$ , while such knowledge is “shunned, ignored, or kept out of mind, and the subject acts in various ways” as if the subject believed that  $p$ , though other behaviour may betray the knowledge that not- $p$  (p. 514); and
- (c) intermediate accounts, where the subject both believes that  $p$  and believes that not- $p$  (Davidson, 1985), or where what the subject believes remains indeterminate (e.g., Funkhouser, 2009).

All these approaches agree on the discrepancy between the attitude toward  $p$  held by the self-deceiver and the attitude the self-deceiver should have, given the available evidence. Lynch adds that it seems to be characteristic of self-deception that the subject encounters the countervailing evidence, while this is not typical of wishful ignorance. For in willful ignorance the subject successfully manages to avoid evidence altogether and does this voluntarily and intentionally. This guarantees that willful ignorance lacks the encounter with evidence that is typical of self-deception.

However, Lynch thinks that this is no conclusive evidence that willful ignorance is not a kind of self-deception after all. For the former could, for example, be a nonparadigmatic case of self-deception. Lynch contemplates this possibility because he is persuaded that, if we had an analysis that did not rely on any of the three views listed above, we could get different results. If such an analysis were available, maybe we could be in a position to see that willful ignorance and self-deception are two of a kind, notwithstanding their obvious differences. I have reasons to think that the theory I will develop could be a promising candidate for being such a view. But before I move on to it, it is important to see why all the paradigmatic-state accounts are inadequate.

All the approaches seen thus far lead us to think that there must be a characteristic, paradigmatic state the self-deceiver is in. They do so by looking statically at the final product of the self-deceptive process. Even if Lynch seems to be more liberal in admitting that more states than are predicted by a paradigmatic-state account could be instantiated by the self-deceiver, he ends up adopting a rhetoric suggestive of the importance of singling out the paradigmatic state. Inso-

far as all these views look statically at the allegedly final product of self-deception, they can be described as snapshot theories: that is, it is as if they take a static, instantaneous picture of the mental state that a self-deceiver is in at a certain time  $t$ , presumably taken to be representative of the central phase of self-deception, and try to unpack its features so as to meet the constraint of tension.

By doing so, however, these views lay themselves open to a quite obvious objection: how should we deal with the empirical discovery that, with regard to the self-deceptive process, self-deceivers are in mental states other than the paradigmatic one? What if these other states are capable of meeting the constraint of tension, too? Obviously, the answer will be that all those accounts suffer from a restrictedness that puts us on the wrong track in our attempts to give a description of the phenomenon, which is rich enough to include other possible, often empirically instantiated, mental products. This is exactly what I think happens if we go beyond the biasing search for a paradigmatic state and head toward a more accurate focus on the self-deceptive process as a whole.

The analysis I will propose in the next section is given over precisely to showing how we should frame our view of self-deception, by considering the process, its moving forces, its dynamics, and all its possible, evolving, mental products. We will see that there is a multitude of highly tensive and unstable mental states that can be instantiated by a self-deceiver, and which have been unjustifiably excluded by paradigmatic-state accounts. However, as long as we persist in trying to freeze self-deception in a certain state instantiated at a certain time  $t$ , we lose the chance to give citizenship to all those other mental states. Insofar as a dynamic view is interested in focusing on the process instead, it leads us to liberalize the varieties of mental states in which a self-deceiver may be. Let me then move on to an outline of such a dynamic, liberal view.

### 3. A DYNAMIC, LIBERAL VIEW OF SELF-DECEPTION

As we have seen, virtually all scholars who study self-deception agree that there are two main forces that set the process in motion—namely, motivation that  $p$  be true, and the thrust of evidence that points to not- $p$ . Both factors are active together, and most likely they create a force field that can be, and in fact often is, highly dynamic. For obviously these factors can vary, and co-vary, depending on contingencies that can occur over time. For instance, there may be times when self-deceived subjects feel more strongly the pull of their motivation that  $p$  be true. Accordingly, they may engage more intensely in their biased treatment of the evidence or of the hypotheses associated with what the evidence suggests. At other times, instead, the encounter with evidence that not- $p$  may be more pressing, either because further evidence is provided or because the evidence already possessed is now seen in a less prejudiced light or else because the motivation that  $p$  be true as such is weakened by other intervening factors that have nothing to do with evidence (e.g., a diminished interest in  $p$  being true on the part of the subject).



There are also presumably noncontingent factors that can intervene in the dynamic of the process. For instance, these may be certain quite stable features of the subject's psychology. For example, if a subject is well trained to treat evidence impartially, even if motivation may have more or less momentarily suspended, shunned, or weakened that epistemic virtue, the latter may at some point just spontaneously strike back and correct, at least for a while, the motivationally distorted treatment of evidence. This may not guarantee the subject's complete exit from self-deception, as it may be the case that the motivation that  $p$  is true strikes back in turn once again. But it can certainly change the specific mental state self-deceiver is in at least temporarily. Perhaps, before such epistemic virtue made its claims felt, the subject placed more confidence in  $p$  than he or she does now. But if the drives of motivation rise forcefully again, the subject may revert once again to a higher degree of confidence in  $p$ .

Let me now expand on the possible mental states a self-deceiver can experience with regard to the self-deceptive process. It may not simply be the case that degrees of confidence vary, as Lynch (2012) correctly diagnoses. It may also be the case that a subject can at times temporarily reach even a state of full-blown belief that  $p$ , while, owing to the variations in the factors described, the same subject can revert to less than that, even to the antipode of believing not- $p$ . Of course, given the dynamics at work, none of these attitudes seems set to last. Or else there may be times when the subject reaches a false second-order belief that he or she has a first-order believe that  $p$ , while truly believing that not- $p$ , as Funkhouser (2005) requires. And there may also be moments when the subject is in an intermediate, indeterminate state of mind, possibly even recognizing it as such.

As we see, the dynamic force field that self-deception amounts to can easily instantiate a vacillation between  $p$  and not- $p$ , one that can (and often does) include a variety of attitudes toward  $p$  and not- $p$ . No attitudes can be excluded in principle, and what temporal extension and qualitative intensity such vacillation may have is a totally empirical question. It all depends on the individual subject, that subject's specificities, and the contingent evolution of his or her practical and epistemic interaction with evidence and reality.

This vacillation over time is best captured, I think, as an "attitude seesaw." There is no reason to exclude from it a priori any of the states that have been indicated as paradigmatic by different scholars, including those that they judge nonparadigmatic, such as escapism. There seems to be no a priori reason why a subject could not at times engage in escapism as well, perhaps emerging from it after a while. Or a subject might just start with escapism and then move on to other less extreme attempts to deal with reality. Nothing in escapism suggests that it cannot be irreversible, nor is there anything in self-deception to suggest that escapism cannot be one of its more-or-less temporary outcomes. The same interpretative line may be applied to willful ignorance, I think. Willful ignorance may well be a phase in the self-deceptive process, earlier or later on—one that can later be

reversed by new variations within the force field, or that can be entered from another state.

Lynch gets close to this when he says that a self-deceiver may at times go as far as self-delusion. He thinks that self-delusion is not a paradigmatic case of self-deception, but he shows an awareness of the variations to which I am drawing attention. My proposal, however, is more radical, and certainly more liberal: there is no need to persist in looking for a paradigmatic state when it is clear that, by its very nature, the process of self-deception can be, and often definitely is, highly dynamic, evolving and varying according to the opposing forces at work in it.

It is apparent that, in a liberal view such as the one I am putting forward, tension is fully preserved. First, such a liberal view preserves the tension that has been associated with a certain single state, or set of states, by paradigmatic-state-account theorists: if and when the state occurs, it has the tension that its proponents correctly attribute to it. In addition, however, there is a further tension that my account guarantees, and it is not clear that paradigmatic-state accounts take it into account: namely, the tension that a more or less prolonged vacillation produces over time.<sup>5</sup> Note that more self-reflective subjects could also experience a sort of metacognitive tension—that is, they find themselves on an attitude seesaw that they can intuit as such. To be sure, however, even if subjects do not reach this metacognitive level of self-reflection, the psychological effect of going on such a seesaw may make them experience tension, presumably of a more opaque kind.<sup>6</sup>

I said that it is a totally empirical question whether a subject enters into any of the possible states that the force field of self-deception allows. Equally, it is a totally empirical question how long such an attitude seesaw can last. Only a case-by-case empirical analysis of specific self-deceptive processes in individual subjects can give an answer to this question.

Let me then briefly recapitulate the main tenets of my proposal.

When one formulates a theory of self-deception, there seems to be no renouncing a couple of necessary constraints—namely, that

- (i) the process is triggered by a motivational state that leads the subject to wish that things stand in a certain, desired way (*p*);  
and
- (ii) the subject driven by such motivation struggles with evidence, or easily accessible evidence, that suggests how things really stand.

The clash between motivation and evidence makes the process highly tense from a psychological point of view. Psychological tension is a descriptive feature of self-deception that virtually all scholars refuse to renounce; rather, they all

visibly want it to be preserved, predicted, and accounted for. When this *desideratum* is combined with the bias in favour of the search for a paradigmatic state, scholars may feel the pressure to look for one single, characteristic state for self-deception, which is also highly tensive and unstable in itself. This combination of drives has led to several competing theories of the characteristic state of mind of self-deception, where what is at stake is the kind of state that can best account for instability, while also satisfying our search for a paradigmatic state of self-deception.

However, we will see that if we subscribe to (i) and (ii), we must accept the consequence that whatever mental state turns out to be compatible with the combined action of (i) and (ii) must be considered characteristic of self-deception. And clearly, there is no reason why we should not extend this consequence to mental states that may turn out to be even quite distant from the paradigmatic one.

Since a wide variety of states are compatible with (i) and (ii), my proposal is that

(A) we should liberalize the types of mental states that are representative of self-deception; and

(B) to avoid remaining on a misleading track, we should replace any paradigmatic-state accounts of self-deception with a dynamic view of the self-deceptive process.

Lynch (2016) provided an argument to show that willful ignorance is not a case of self-deception, whether paradigmatic or nonparadigmatic; therefore, it is a different kind of phenomenon. I think, however, that if we liberalize self-deception along the lines suggested, we are in a position to see that it may be the case that willful ignorance is sometimes a phase along the process of self-deception. Willful ignorance may well be the kind of phenomenon that Lynch diagnoses it as, and may fully retain its characteristic. The two phenomena need not be conflated. Yet willful ignorance could surface along the liberalized self-deceptive process as a possible stage of it. It is also important to note that my liberal view of the self-deceptive process does not rely on any of the three views of self-deception that, according to Lynch, could be used to distinguish willful ignorance from self-deception. Since I am not relying on any of the three, I am in a position to include willful ignorance as a possible stage of self-deception, and to do so by Lynch's own standards. For I am not proposing an unwarranted-belief account, where the subject ends up believing that *p* self-deceptively (Mele, 1997, is an example of such an account). I am not proposing an implicit-knowledge account either (e.g., Bach, 1981), where the subject does not believe that *p*, but recognizes the truth of not-*p*, while that knowledge is "shunned, ignored, or kept out of mind," and the subject's behaviour suggests that he or she believes that *p*, though other behaviour may betray the knowledge that not-*p* (p. 514). Nor do I propose an intermediate account, where the subject both believes that *p* and believes that not-*p* (Davidson, 1985), or where what the subject believes remains indeterminate (e.g., Funkhouser, 2009). Rather, I am proposing a view

whose focus is on a process within which more than is dreamed of by our paradigmatic-state-account philosophy of self-deception may happen.

My sense is that there might be room to apply a liberal solution to twisted self-deception as well. Twist-self-deception (Mele, 1999) is typically described as a case of self-deception where a subject does not end up believing what he or she desires or wants to be true. Rather, the subject ends up believing what he or she fears and, in any case, does *not* want to be true. Even if investigating this would lead me too far from my present purposes, it seems credible that a twisted self-deceiver may experience an attitude seesaw fully compatible with (i) and (ii). If this is correct, then twisted self-deception should cease to be seen as a nonparadigmatic case of self-deception, as, Lynch notes, it is still generally considered (2016, p. 513).

One may wonder whether any liberal view is too liberal after all. That is to say, does a liberal view risk being too broad, thus erring on the side of overinclusiveness? Were a liberal view to include more states than really fall within the self-deceptive kind, we would lose the unity of the phenomenon, as well as its specificity.

I do not think that overinclusiveness is a genuine risk for a liberal view, except when liberality is completely unrestricted. But I set (i) and (ii) as constraints for self-deception. Thus, I have reason to think that as long as (i) and (ii) are active, none of the states that can possibly be entered by subjects during their self-deceptive attitude seesaw is outside the phenomenon of self-deception. Of course, should either (i) or (ii) stop being active, then the subjects might well enter another kind of phenomenon altogether, such as permanent self-delusion or, in the event of their exiting self-deception completely, full adherence to reality.

I conclude with a final remark on what my analysis seems to suggest in terms of the adequacy of the conceptual apparatus we deploy for capturing self-deception. Even if I agree with Lynch (2012) that we should ameliorate our conceptual categories and look for more fine-grained concepts to capture psychological reality (p. 438), I also think that we should be ready to apply the categories we already have more liberally whenever our psychology shows a complexity that no single traditional mental category can capture in isolation. Sometimes, psychological complexity just requires us not to force and freeze it into snapshots. Rather it clearly invites us to look more closely at the richness embedded both in its static dimension and in its temporally evolving dynamic.

## ACKNOWLEDGMENTS

I am grateful to the audience of the Latin Conference of Analytic Philosophy held in Paris at the Jean Nicod Institute in July, 2013, for helpful questions on an early outline of this argument.

I am also indebted to two anonymous reviewers of this journal for their valuable comments on an early version of this paper.

## NOTES

- <sup>1</sup> Famously, the other paradox is the “dynamic paradox.” Since it is not immediately useful for my argument, although it is closely connected to it, I will not go into it in this paper.
- <sup>2</sup> I will not address the partitioning solutions to the static paradox here. See Deweese-Boyd (2017).
- <sup>3</sup> I thank one anonymous reviewer of this journal for pointing this extremely crucial aspect out to me.
- <sup>4</sup> Funkhouser (2005) relies on Nelkin (2002) here. They both think that self-deceiver reaches a false, second-order belief that he or she believes that  $p$  because that self-deceiver is primarily moved by a mind-directed desire to believe that  $p$ , not by a world-directed desire that  $p$  be true. For views that go in the same direction, see also Holton (2001) and Fernández (2013). For comments on the mind-directed desire, see also Pedrini (2012; 2013).
- <sup>5</sup> I am grateful to one anonymous reviewer of this journal for pointing out the importance of emphasizing this variety of tension.
- <sup>6</sup> Cf. Pedrini (2018).

## REFERENCES

- Bach, Kent, "An Analysis of Self-Deception," *Philosophy and Phenomenological Research*, vol. 41, 1981, p. 351–370.
- Bilgrami, Akeel, *Self-Knowledge and Resentment*, Cambridge, MA, Harvard University Press, 2006.
- Davidson, Davidson, "Deception and Division," in *Actions and Events*, E. LePore and B. McLaughlin (eds.), New York, Basil Blackwell, 1985.
- Deweese-Boyd, Ian, "Self-Deception," *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), Edward N. Zalta (ed.), URL: <https://plato.stanford.edu/archives/fall2017/entries/self-deception/>
- Egan, Andy, "Imagination, Delusion, and Self-Deception," in Bayne, T. and J. Fernández (eds.), *Delusion, Self-Deception, and Affective Influences on Belief-Formation*, Hove, Sussex, Psychology Press, 2009.
- Edwards, Sophie, "Nondoxasticism about Self-Deception," *Dialectica*, vol. 67, no. 3, 2013, p. 265–282.
- Fernández, Jordi, "Self-Deception and Self-Knowledge," *Philosophical Studies*, vol. 162, no. 2, 2013, p. 379–400.
- Funkhouser, Eric, "Do the Self-Deceived Get What They Want?" *Pacific Philosophical Quarterly*, vol. 86, no. 3, 2005, p. 295–312.
- , "Self-Deception and the Limits of Folk Psychology," *Social Theory and Practice*, vol. 35, no. 1, 2009, p. 1–13.
- Gendler, Tamara S., "Self-Deception as Pretense," *Philosophical Perspectives*, vol. 21, no. 1, 2007, p. 231–258.
- Holton, Richard, "What is the Role of the Self in Self-Deception?" *Proceedings of the Aristotelian Society*, vol. 101, no. 1, 2001, p. 53–69.
- Lazar, Ariela, "Deceiving Oneself or Self-Deceived?" *Mind*, vol. 108, no. 430, 1999, p. 263–290.
- Longeway, John, "The Rationality of Self-Deception and Escapism," *Behavior and Philosophy*, vol. 18, no. 2, 1990, p. 1–19.
- Lynch, Kevin, "On the 'Tension' Inherent in Self-Deception," *Philosophical Psychology*, vol. 25, no. 3, 2012, p. 433–450.
- , "Willful Ignorance and Self-Deception," *Philosophical Studies*, vol. 173, no. 2, 2016, p. 505–523.
- Mele, Alfred, "Real Self-Deception," *Behavioral and Brain Sciences*, vol. 20, no. 1, 1997, p. 91–102.
- , "Twisted Self-Deception," *Philosophical Psychology*, vol. 12, no. 2, 1999, p. 117–137.
- , *Self-Deception Unmasked*, Princeton, Princeton University Press, 2001.

Nelkin, Dana K., "Self-Deception, Motivation, and the Desire To Believe," *Pacific Philosophical Quarterly*, vol. 83, no. 4, 2002, p. 384–406.

Pedrini, Patrizia, "What Does the Self-Deceiver Want?" *Humana.Mente: Journal of Philosophical Studies*, Pedrini, Patrizia (ed.), vol. 5, no. 20, 2012, p. 141-157.

———, *L'autoinganno. Che cos'è e come funziona*, Rome, Laterza, 2013.

———, "The 'Crux' of Internal Promptings," in Patrizia Pedrini and Julie Kirsch (eds.), *Third-Person Self-Knowledge, Self-Interpretation, and Narrative*, Springer, 2018.