



Political Self-Deception revisited: reply to comments

Anna Elisabetta Galeotti

To cite this article: Anna Elisabetta Galeotti (2020) Political Self-Deception revisited: reply to comments, *Ethics & Global Politics*, 13:4, 56-69

To link to this article: <https://doi.org/10.1080/16544951.2020.1846947>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 10 Dec 2020.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

ARTICLE

 OPEN ACCESS  Check for updates

Political Self-Deception revisited: reply to comments

Anna Elisabetta Galeotti 

Dept. of Humanities, University of Piemonte Orientale, Vercelli, Italy

ABSTRACT

The article replies to the five comments to *Political Self-Deception*, from the more philosophical and epistemic remarks to the more political and historical ones. In the end, it summarizes the main points of the book as suggested by the discussion with the five comments.

ARTICLE HISTORY

Received 14 August 2020
Revised 2 November 2020
Accepted 3 November 2020

KEYWORDS

Self-Deception; invisible hand; epistemic vice; evidence; groupthink; explanatory power

Introductory note

The five comments to my *Political Self-Deception* (2018) provide original insights on my work from different viewpoints; in this sense, they are complementary to each other and to my study: besides raising critical remarks, each of them prospects a different line of inquiry, and wonders whether the core idea of my book could be developed and applied in that area. From their different perspectives, each contribution enlarges my work and prepares the ground for specific researches and applications in different directions and disciplines, and I am grateful for their insightful criticisms. I shall consider each contribution in turn, starting from the comments more focused on the philosophical and epistemic aspects of SD, corresponding to the first part of my work, and then going to the more political and historical perspectives tackling with themes that I take up in the second part of the book.

Self-Deception or epistemic vices?

Neil Manson's 'Political Self-Deception and Epistemic Vices' expresses his philosophical concerns, questioning, on the one side, the use of the invisible hand model to explain SD, and, on the other, the explanations of political failures by means of SD instead of epistemic vices. Manson underlines a difference in the invisible hand explanations in economic science compared to my SD explanation, and such difference makes him doubt that what I employ is indeed an invisible hand model. In economic explanations, individual actions to certain individual ends compose themselves in an unintended outcome, serving some different and unexpected collective purpose (Nozick 1974, 1977). By contrast, 'the hand, in individual SD, does not seem to be very much guided

CONTACT Anna Elisabetta Galeotti  annaelisabettagaleotti@gmail.com  Dept. of Humanities, University of Piemonte Orientale, Vercelli, Italy

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

by the individual's own motives'; it seems rather a hand working 'outside the agent's consciousness'. Manson's remark gives me the opportunity to specify the invisible hand model, whose role in my argument is very important for it allows me to keep together the intentionality of SD process with the unintentionality of SD outcome (the deceptive belief) (*Political Self-Deception* p.41). In fact, I hold that the steps undertaken by the agent to check whether p is true are intentionally taken and are motivated by the attempt to dispel the anxiety created by the evidence contrary to p . Thus, contrary to what Manson has remarked, the hand is indeed guided by the individual's own motives. For the agent openly engages in the process of thinking and ruminating with the explicit motivation to scrutinize the evidence, but then, led astray by her wish that p be true, is exposed to biases and faulty twists, ending up with the desired yet unintentionally brought about belief that p is true. Agents are in fact unaware of their biased thinking, as much as social agents are unaware of the supply-demand law. In the social science, agents are unaware of the overall effect of the composition of their own actions; similarly, in SD, the agent is unaware of the overall effect of the single steps she is taking in her reasoning and of their faultiness. In either case, actions (or steps) elsewhere directed jointly bring about an unintended, yet beneficial outcome. In case of SD, the outcome can be said beneficial insofar as it corresponds to the agent's wish and eases her anxiety, at least in the short term. In brief, it seems to me that the invisible hand model neatly fits the process leading to SD providing an account of the three elements for SD production: the intentional steps, the causal biasing, and the unintended deceptive belief that results from the process.

The second critical remarks by Manson refer to the area of study of epistemic virtues and vices (Zagzebski 1996; Fairweather and Zagzebski 2001; Cassam 2019) which, to his mind, can provide a more nuanced explanation of the various epistemic failings leading to bad political decisions. Arrogance, inattentiveness, failure to give credibility to certain sources based on prejudice, to his mind, can offer an idiographic explanation of certain mistakes, whereas SD provides only a general explanation, but with much less details. Manson himself responds to his remark saying that the two explanations can supplement each other, the first providing a sort of general model, while the second getting into the specifics of singular cases. I doubt that Manson's proposed solution would work, given that the epistemic vice perspective, compared with SD, does not simply provide a more detailed and nuanced tool for the analysis of epistemic failings; the epistemic vice approach views mistakes as the consequences of bad character dispositions, and of the agent's failure to cultivate good epistemic habits. Cassam (2019) even understands cognitive biases as the possible outcome of epistemic vices. I tend to resist Cassam's suggestion for cognitive biases seems to be spread among all sorts of people, even intelligent and generally accurate. In general, as I have argued at length in the book, I do not think that SD depends on a certain type of character (corrupted by epistemic vices); on the basis of a wide literature in experimental psychology, SD seems rather to be produced under the effect of given typical circumstances. Anyone can be and is sometimes struck by SD under certain circumstances, though I can agree that some epistemic vices make it easier to fall prey of SD. That is why my suggestions at SD prevention mainly rely on institutional mechanisms rather than on education in epistemic virtues. Most importantly, as Manson himself acknowledges, SD points to the motivation behind the faulty reasoning, which is instead absent

in the epistemic vice explanation, and explains the mistake as the outcome of motivated reasoning. In case, the epistemic vice explanation seems more appropriate for simple mistakes than for proper cases of motivated twisted reasoning. Finally, I actually try to provide a general model to explain not all faulty reasoning and decisions, but the small portion where a motivation can be detected as lurking behind the process. In other words, my problem is not epistemic failing per se, but rather motivated reasoning against the available evidence. In sum, my reservation on Manson's suggestion to supplement the general model of SD with the idiographic tools from the epistemic vice literature refers to two main reasons. On the one hand, SD is not a general model to interpret cognitive mistakes, but rather an account of motivated reasoning, and, on the other, my account of SD does not imply a link with a certain defective character, as it is the case with epistemic vices.

The available evidence and the impartial umpire

I shall now consider the comments by Alfred Moore 'Reality Check: Can Impartial Umpires Solve the Problem of Political Self Deception?', for his contribution too is focused on epistemic matters. More precisely, the issue he raises concerns the evidence that constitutes the benchmark for detecting self-deception. In my own definition 'SD is believing that P, under the influence of the desire that P be the case, when *the available evidence* would lead an unmotivated cognizer to conclude that $\sim P$ ' (p.19). In order for SD to be detected, then the available evidence is paramount as well as some unmotivated observer who plays the role of an umpire, as Moore says. Moore points out that the available evidence is rarely clear-cut, but, especially in politics, usually complex, fragmented, available only in pieces and bits, often contradictory. And he is right: often only in hindsight one can have a clear view of the data. That is also the reason why singling out cases of SD in politics is hard, and there is always a controversy about what evidence was actually and clearly available at the time of the decision. Moore's remark, however, goes deeper than that: not only is the available evidence often foggy and complex, but, moreover, it is not an 'impartial umpire' who, by virtue of his impartiality, can appraise it correctly. The reason he offers is twofold: first, the assessing of evidence always implies value-laden judgements, hence the reference to a dispassionate judge is misplaced; second, the relevance of disagreement and conflict in reaching collective decision is not given an adequate role in my suggestion of prophylactic measures against SD. Moore's proposal is instead to correct the misperception brought about by SD with other situated and partial viewpoints. The plurality of misperceptions would induce conflict and disagreement, and out of that, a collective process will develop where individual self-deceivers will somehow correct each other and reach a better epistemic result, as predicted by studies such as Sperber and Mercier (2011) and Landemore (2012). In spelling out his proposal, Moore criticizes also the ambivalence he sees in my presentation of SD as a collective process. In Moore's view, I describe collective SD halfway between a summative and an 'emergentist' (supervenient) process. If it is just summative, then the collective dimension is actually downplayed, while he thinks that it should be placed in the forefront. In general, Moore stresses that the cognitive distortions and defective decisions induced by SD should not be fought with the attempt to reinstate a correct epistemic process via an impartial, dispassionate observer, given that all appraisals of evidence and data interpretations always embody value-judgements and perspectival

approaches. Consequently, his suggestion is to let disagreement and conflict develop towards a more balanced and sound outcome, as the result of collective reasoning. In other words, the collective reasoning, seen as an emergentist process, could be the treatment of the cognitive distortions, whether brought about by SD or by the different perspectival positions of each individual in the process.

I shall try to respond, starting with the issue of the ‘dispassionate umpire’ versus ‘diverse participants’. Moore rightly points out that evidence assessment is hardly a value-free, impartial process; yet, one thing is to assess the evidence from a situated point of view, and another to have a specific motivation for considering the evidence in a biased way, namely that the wish that p be true. I am not advocating a perfect impartiality and detachment from values and emotions; yet, the fact that our evaluations are always situated and perspectival does not imply that we cannot be free from the influence of the emotionally overloaded wish that p in our reasoning. By ‘dispassionate’ I precisely mean not sharing the ‘rebellious’ wish that p , which is what is relevant for SD process. To rephrase my own words, the cognizer is not unmotivated in general, but is free from the motivation with reference to p . What happens with cases of personal SD is precisely that the network of friends and relatives is able to detect the self-deception, given that none of them shares the same anxious wish that p with the subject. Hence, the subject’s belief looks aligned with the desire that p rather than with facts that the observers find more pending in favour of $\sim p$.

Moving now to political SD, the context is much more complicated and messier than personal life, nevertheless the problem is similar, given that SD can be detected if data are approached with eyes not clouded by the *specific wish that p be the case*. Having the eyes unclouded by the wish that p does not imply to have a perfectly clear eyesight. Again, the problem is to break the vicious circle between the wish that p and the subsequent biased search and biased treatment of data in favour of p . I agree that a perfectly unbiased treatment of data is out of question, but what is relevant to detect SD is the absence of the influence of the wish that p . One thing is that the data are approached from a given perspective, another is that the wish that p sets in motion a process that results in the *counter-evidential belief that p* . In this respect SD cannot be confused with the perspectival nature of knowledge, for SD implies something more specific, namely that the belief that p runs *against* the available evidence, because of the motivation that p . The available evidence may not be beyond dispute, for sure, yet what is precisely counter-evidential is relatively easier to establish.

Let us now come to collective self-deception and to Moore’s suggestion that individual SD can be a weapon against collective SD. I have actually described the process of collective SD at some length, as I stated: ‘contrary to what happens in cases of personal SD, political SD is almost always a collective product’ (p.100). The collective product is the result of the positions of different individuals and of *the power structure of the group*.¹ In theory, a group provides more possibilities to confront information and data critically, to check biases and idiosyncrasies, to test hypothesis and to get to a better decision, but, in fact, the group can harbour pathologies depending on its internal power structure and on its ideological connotation (Janis 1982). Precisely the studies,

¹The stress on the power structure of the group is the point where my position of collective SD differs from Janis’s Groupthink (Janis 1982).

quoted by Moore, on the outperformance of diverse groups over intelligent individuals' show that the epistemic superiority of diverse groups works if anyone shares the problem-solving approach with the paramount motivation to reach the best result (Landemore 2012). SD, however, has precisely to do with motivations interfering with reasoning. As I have described, the members of the group need not share the same motivation, but the power structure and the fear to displease the leader and fall off his or her favour affect everyone. The different steps leading to SD are taken by different agents; for example, the conviction that p be true by some is taken by others as evidence that p be true. At the end, there is a process of reciprocal reinforcement of the deceptive belief so that it becomes collectively shared. In this case, collective reasoning has moreover the perverse effect of inducing group members to consider the stakes of the collective decision less than his or her personal stakes within the group, and that, in turn, lowers the costs for inaccuracy for each and helps the self-deceptive process to start. Thus, coming back to Moore's remarks, collective SD is produced by the single steps taken by each individual member, but the outcome, collective SD, exhibits some supervenient properties, such as the illusion of invulnerability. Contrary to Moore's view, though, I do not think that collective SD and individual SD can work against each other, providing the required check on decision making. In my conjectural reconstruction, the SD of individual members and of the whole group goes in the same direction; and even if in the group someone harbours doubts, hardly such doubts will be voiced effectively so as to check the decision process, because of the power structure and on the dependency on the leader. To put it differently, contrary to what Moore seems to think, the collective process is not driven by the search for accuracy and best outcome, as it is the collective reasoning by the groups considered by Mercier and Landemore, but is motivated as well as individual SD. I have especially stressed the power structure as the locus of epistemic distortions, and that is why, with reference to prophylactic measures, I am afraid that a devil's advocate may not be effective for he or she may be taken in by the group structure. The quest for the independence of overseers concerns precisely the freedom from the power structure, and from the identification with the leader's wish, and not the independence of an undetached point of view. My impression is that Moore's suggestion to fight collective SD with individual SD equates SD with a perspectival approach to data, so that confronting many diverse perspectival views, a more balanced opinion may emerge. Yet SD is not simply a value-laden, oriented consideration of evidence: it is a *specifically* biased consideration of evidence, due to the influence of the emotionally loaded wish that p , resulting not just in a 'partial' or 'oriented' view, but precisely in a *counter-evidential* belief. Thus while I agree with Moore's stress on the oriented, partial and perspectival consideration of data, I like to keep SD as something more specific than such a general character of human knowledge.

Political theory and empirical evidence

Alice Baderin's comments too, in 'From Political SD to SD in Political Theory', are centred on the issue of evidence, though in a different way than Alfred Moore's. More precisely, she asks about the room for empirical evidence in political theorizing, exploring the possibility of extending the analysis of SD from politics to political theory. She points out that many political theorists may be found engaged in motivated

reasoning when they try to match their normative ideals with empirical evidence. The latter is often considered in a biased way, specifically tailored in order to be fitting their normative positions. In fact, as Baderin acknowledges, this is more a case of believing *beyond evidence* than believing *against evidence*, a case characterizing wishful thinking more than SD. Similarly, she notes that the wish to be edgy and provocative often underlies theoretical enterprises built on disagreement, where theory is guided and moulded out of the motivation to take a controversial position that no one can just ignore. Again, it is not clear that such is a case of SD or simply of letting theory be led by extra-theoretical desires. In general, I think that Baderin's attempt to apply motivated reasoning at the meta-theoretical level is very original and revealing, and that the avoidance of similar pitfalls concerning the role of empirical evidence in political theory should be taught to students as part of a sound methodology.

With reference to the role of empirical evidence, Baderin also raises a critical remark to my political SD, picking out what she holds to be a weakness in the empirical foundation of my normative claims. I hold that the favourable circumstances for political SD to take place are exceptional crises calling for momentous decisions under pressure of time. From these circumstances, my argument moves to the prophylactic measures, for, thanks to the detection of favourable circumstances, SD is at least partially predictable in advance, while neither lies nor mistakes can be foreseen *ex ante*. On this argument, Baderin advances two critical comments, the first concerning the empirical circumstances that ground the normative claim about prophylactic measures, the second concerning the consideration of lying as unpredictable. Relative to the first point, Baderin precisely refers to time pressure and dramatic decision making. She observes that such circumstances are absent not only in cases of personal SD, but also in political SD when, for example, politicians are covering up past mistakes. How can they be the typical circumstance for SD to arise? Moreover, she stresses a sort of incongruence between the exceptionality of circumstances and the predictability of the subsequent SD. I shall immediately clear this latter remark: exceptional circumstances, as a rule, are not anticipated, but the response to exceptional circumstances can be, as all disaster management studies testify. In our case, SD is the (at least partially) predictable response to exceptional, hence unanticipated, circumstances. Thus, there is no incongruence here. Focusing now on her first remark, Baderin is right in noting that time-pressure and momentous decision-making are not typical either of personal SD or of political SD, when at issue there is the 'sour grapes' mechanism for the cognitive dissonance reduction of past mistakes (Elster 1983). However, both time pressure and momentous decision making are simply instantiations of the contextual condition for SD to start that I have presented both at the end of chapter 1 (p.56) and at the beginning of chapter 3 (p.80). The contextual condition varies according to the kind and type of SD, as long as it is something that arouses anxiety in the subject and emotionally overloads the wish that *p*. In personal cases, the wish that *p*, which is something very important for the subject's well-being, is emotionally turned on by the mere suspicion of evidence contrary to *p*. In the context of international politics, the wish that *p* is not connected to the personal well-being of the decision-maker, but rather to her strategic goals or to her popular image. In that case, the wish that *p* becomes anxious under the exceptional circumstances and under evidence unfavourable to *p*, which leads the decision maker towards a biased search for reconfirming *p*.

What about then the sour-grapes type of political SD (pp.93–95)? In that case, the wish to defend one's cherished image against a failure is anxious and does not need time pressure or exceptional decision to become emotionally overloaded. The painful knowledge of one's failure is sufficient to push the subject to reassess the criterial evidence and realign what happened with his self-image. In sum, the contextual element, no matter what form it may take, has always the role of turning on the crucial wish leading to the (biased) appraisal of data and the search for a solution in line with the wish that p , despite the available evidence. True, the cases that I have analysed in depth in the second part of the book have all involved decision-making under conditions of time pressure and momentous crises, which in such cases represented the contextual element favourable to set in motion an SD process. And yet, time pressure and momentous decisions are not the only contextual condition arousing the wish that p and leading to a biased search of data.

Finally, Baderin contests my point that only SD can be predictable, while lies and mistakes are not. More precisely, she focuses on lying, and on the growing body of studies engaged in finding conditions under which lying may more easily take place, or may more easily be detectable in advance. Actually, it is not clear whether such conditions make lying predictable or just detectable. Certainly, the study of political speech pattern for lie detector is not useful for predicting lying (Bond et al. 2019), but also the list of conditions that Baderin gets from behavioural economics, which are tailored on individuals and not on groups, says more about how to discover a lie than how to predict it.² Thus, while I cannot rule out the possibility of predicting lying in general, it seems to me that political SD, which is a collective product, is more clearly predictable than lying, which is an individual act.

The realist objection rephrased

Lior Erez, in his 'What is Political about Political Self-Deception?', raises the question which may have more disruptive effects on my work, for he asks from a realist perspective whether SD may represent a useful category in political analysis. I have actually considered such a question myself, in the Introduction, under the label 'the realist objection', and concluded that even from a realist viewpoint, SD can be useful for its explanatory role in making sense of certain political failure. Erez rebuts my argument. He starts by saying that realists have no difficulty in acknowledging that politics, like any other social sphere, has its own share of SD. The point is however whether political SD affects political judgement, in the double sense of negatively influencing political decisions and of inducing a negative evaluation of the policy chosen via a self-deceptive process. If that were the case, then SD would be a relevant category in political analysis, as held in my book. But Erez answers in the negative in either case, for to his mind neither SD necessarily entails poor political judgement, nor the political analyst has much to criticize the self-deceiver. The Realist is interested in outcomes and consequences, as Erez reminds us, and whether the political agent is lying or is self-

²She quotes: tiredness and cognitive depletion; when the payoff is closer to the decision-point; awareness of future altruistic opportunities; a feeling of anonymity at the decision-point; seeking to cover a loss, rather than to make a gain; and when people with whom you identify are lying (Jacobsen, Fosgaard, and Pascual-Ezama 2018).

deceived does not change the political scientist's evaluation of decisions, actions and policies, as long as they are effective, given that the realist does not think that normative considerations are relevant for political judgements. To be fair: Erez thinks that there is room for political normativity, only well marked off and distinct from moral normativity, and stresses that private morality and political action are two separate spheres. In passing, no normative political theorist imagines that public reasons and public ethics should be the simple extension of private morality (Hampshire 1978; Larmore 2020). What are the specific normative questions appropriate to politics are left unclear, and I suppose that they have to do with effectiveness and efficiency relative to ends whose worth may be considered only in terms of feasibility. With this portrait of the contemporary realist, Erez considers my response to the Realist Objection ineffective for he denies that being self-deceived, that is believing something contrary to data, implies to bring about poor policies. Political actors, in his view, are not 'neutral assessors of evidence and probability in the political landscape' for they actively shape that landscape by their decisions and actions. Hence, the evidence is not 'out there', but rather created and moulded by human action. Now, no one denies that human action is causally efficacious in changing the state of things, but this trivial fact does not entail that in the decision moment the agent does not face a given evidence. The evidence may be complex and difficult to decipher, but it is nevertheless the given context from which decisions are taken. It seems to me that there is a non-sequitur in Erez's argument: the fact that politicians are active in the context they are immersed in and contribute to change with their actions does not imply that their actions can dispense with a proper appraisal of the context, as if decisions take place in a void. If I say that the invasion of Iraq is required by the presence of WMD, and after the invasion the WMD are not found, despite the efforts and costs spent in several searches, then I have manifestly and publicly failed the invasion, or better I failed the reasons why the invasion had been justified internationally. Now if I had lied about that, I might have avoided ending up so exposed. Either I could have placed some weapons there to be found, or at least I would have avoided the searches. Thus, it seems to me that being self-deceived about certain facts does not fare good in terms of success and good result. It is true that sometimes a leap of faith may bring about the desirable consequences (the well-known self-fulfilling prophecy) and that excessive prudence may sometimes be detrimental. Yet, first, I will underline the 'sometimes', and second, one thing is the leap of faith, and another is going against hard facts. I concede that what the hard facts are is often difficult to establish in political reality, but precisely when the evidence is weighting in one direction decidedly, ignoring it and let oneself be led by one's wishes is highly risky (and with the lives of other people) and never promising. In case of SD there is not only political responsibility for one's mistakes, which as Erez emphasizes always kicks in for politicians; in such cases there is the moral and epistemic responsibility to have let one's wishes (ends, convictions) distort the judgement. In sum, I do not find Erez's rejoinder to my response to the Realist objection convincing. The fact that the politician is not a theorist or an expert whose task is just to provide data's interpretation, and that he or she has to make decisions and to act does not imply that acting on the grounds of false data is a politically intelligent move. Even if occasionally, rarely indeed, the decision may not turn out in a fiasco, it is more a matter of luck than of political wisdom.

In the second part of his comments, Erez takes up the normative evaluation of SD, as I have discussed in connection with the Iraqi invasion, and considers my two criticisms to the decision-making process, criticisms that he labels as the epistemic objection to SD and the democracy legitimacy objection. The epistemic objection, that is the tendency of collective self-deception to dismiss dissenting voice, is rashly considered in a sceptical fashion, because certainty is not part of political reality. Actually, certainty in the fallibilist epistemology that I favour is not part of human knowledge in general, but that does not exclude that listening to dissenting voice and considering alternative scenarios is an epistemic (and political) virtue as Neil Manson has stressed in his contribution. The second point, the democratic legitimacy objection, concerns not so much SD as the context in which the specific SD about WMD developed. The context was precisely the gap between the selling reasons and the real reasons for the invasion. Here I think that the distance between my position and Erez's depends on the different evaluation of normative questions from the perspective of a normative political theorist and from that of a realist. It does not say much about my political SD, but rather about a fundamental disagreement related to the two approaches. I think that the existence of such a gap from the viewpoint of democratic ethics is extremely problematic because citizens should be informed about the reasons for waging a war and have a right to make up their mind about that, for war is waged by the whole country and there is a collective responsibility about that. Thus, I do not see that the fact of the professional army in any way justifies the disinformation of the public, and dispense with citizens' collective responsibility. Nor do I share the point made by Erez that 'there is no reason to think that the best justifications for engaging in a war are also going to be acceptable public reasons for engaging in a war'. In my perspective, the fact that the decision to waging a war is solidly grounded on a (partisan) interpretation of the national interest and has good chances of success does not provide a sufficient justification if it is not publicly acceptable. In a democracy there is no national interest beyond and above citizens' convictions. I understand that for a realist the only relevant justification is not affected by citizenry acceptability, but that does not satisfy the normative political thinker who does not claim implausible standards, only that such a momentous decision is not taken by deceiving the public. And in the specific case of Iraq, the gap between the real reasons and the selling reasons was in no way justified by the necessity of surprise or similar lame excuses. In cases where there is no possibility to reconcile the effective decision to wage a war with the public reasons for it, maybe there is no cause for war at all.

SD in historical perspective

Shaul Mitelpunkt's comments too are focused on war and international relations, yet from the perspective of the historian and not of the realist political theorist. While the political theorist is interested not only in analysing politics in an unprejudiced way outside the lens of moralized considerations, but also in providing guidelines for efficacious political action in similar circumstances, the historian is interested in reconstructing the broad picture, the general trend within which specific decisions, in case self-deceptive decisions, take place and make sense. In such a perspective, Mitelpunkt acknowledges that the category of SD can be analytically useful to the historian, widening the view of multiple causality with psychological and epistemic components. Nevertheless, Mitelpunkt is worried that my analysis of international crisis

and subsequent failures in terms of a deviant decision making affected by SD suggests the idea that problematic foreign policies are caused by singular wrong decisions. And the idea of prophylactic measures that I propose seems to move along the same direction, namely that once mended SD, then mistaken and wrongful policies, such as the Bay of Pigs, the Vietnam War and the Iraqi invasion would be avoided. Mitelpunkt is right in insisting that the cases that I have considered in *Political Self-Deception* are part of a general trend in American foreign policy that is problematic in its ideological premises and goals. I do not claim that my analysis provides a complete explanation of those events, and I hope to have made that clear in the book. I have also tried to show how shared ideological convictions, such as American exceptionalism and Cold War basic tenets, provided a fertile background for illusions and cognitive distortions at any turning point of those trends. My point is in a way more modest: I have tried to detect specific episodes of SD guiding decision making, when in the confusion of data, commitments and muscular postures, it has been possible to single out specific decisions based on counter-evidential belief. Such decisions were certainly backed by a complexity of factors, rooted in a long-lasting attitude in American foreign policy; nevertheless, they led to specific failures and had long-lasting consequences on that policy. I am not suggesting that, absent SD, American policy concerning Cuba, Vietnam and Iraq, would have adjusted and found a right path. I am only saying that those decisions might possibly be avoided, with all the surrounding deceptions, covering up, secrecy and drama. The treatment of SD – supposing now for a moment that it is an easy thing to do and to make it acceptable – will not mend a poor ideology and will not provide the background for fair convictions, but it may avoid certain specific tragic choices. In this respect, if, as Mitelpunkt emphasizes, the interpretation of a policy as successful may change according to the time span considered, the failure of a policy can more easily be measured by the destruction and death tolls it brings along. Finally, Mitelpunkt repeatedly notes that my analysis does not take into account the issue of democratic SD, or of the SD of the public. This is certainly a very important issue, and possibly more important than what I have been concerned with so far, but I had to start somewhere and the perspective on leaders' decision making has allowed me to reconstruct SD more precisely. Moving towards the people's SD can be the next project.

Concluding remarks

As a conclusion, I shall add some general remarks to the comments that my book has received. In the first place, I want to stress the specific role I assign to SD in politics. This role is not that of replacing either deception or cold mistakes in political decision making. Both deception and mistakes are present and have consequences on policies affecting their success or failure, and affecting the relation with the democratic public. Sometimes, however, mistakes are the result of motivated reasoning, on the one hand, while, on the other, the deception of the public results from the self-deception of politicians. Such cases are precisely cases of SD, where the decisions are informed by beliefs induced by motivation and are counter-evidential, hence false.

One of the examples discussed in the book refers to the Gulf of Tonkin incident, in 1964, during the Vietnam conflict, which never took place, but which was believed to

have occurred because it represented the much desired opportunity for the escalation of the American involvement in Vietnam. As I have reconstructed in detail, the decision to air reprisal against North Vietnam and its alleged attack in the Tonkin Gulf appeared at first to have been beneficial for the Johnson administration in terms of popularity and national support (176–177); and the short-term beneficial effect actually characterizes most cases of SD, both at the personal and at the political level.³ Yet, it was certainly not beneficial for the outcome of the conflict, nor in the end for the decision-makers. Later, in fact, the truth about the Tonkin affair started to surface, with the hearings promoted by Senator Fulbright in 1968, and that finally decided the end of the career of McNamara and of President Johnson, while the war went from bad to worst (MacNamara 1995). Considering this episode, it is hard to see how the harsh decision for the air strike, informed by a deceptive belief, could ever turn out to be a success, as Lior Erez maintains that decisions based on SD can. Similarly, it is difficult to share his view that from a realist perspective, the effect of SD on decision making is irrelevant. Possibly, the realist does not consider that the difference between simply cold mistakes and mistakes produced by SD is relevant, for in both cases, very likely, the decision leads to unwise policies. The explanation of the unwise policies, however, changes in either case: if it is just an epistemic mistake, possibly the reference to epistemic vices, as suggested by Moore, can provide an account. If instead the mistake is motivated, then SD can provide an explanation that, moreover, accounts also the deception of the public resulting from the SD of the leaders. I think that possibly Erez has conflated SD with positive thinking, which is an attitude of optimistic overconfidence pushing people to take bigger risks, sometimes succeeding in hazardous operation. But SD is more specific than that for it implies a belief that is brought about against the available evidence, hence, most of the time, false.

In sum, I hold that SD can display a strong explanatory power, yet only in a limited set of decision-making cases, namely those where the evidence available at the time runs contrary to the conviction of the decision makers. In other words, SD as an analytical category in politics cannot be overused. In that respect, Mitelpunkt's concerns are unfounded. Even when SD can properly make sense of given episodes of decision making, in any case, the explanation in terms of SD requires a larger account concerning why certain motivations were so powerful, and why they came to be emotionally loaded in such a way that led to a biased search for evidence. Such larger account is precisely what historians do when confronting with certain episodes, providing a more complex interpretation reaching further back than the SD explanation of a single decision-making process can ever do. Thus, Mitelpunkt need not worrying about reducing historical interpretations to psychological mechanisms, for political SD can only provide the account of a single step in a very complex process that requires an independent explanation. As said, I am not claiming that SD as a political analytical category can sort out all cases of decision making when momentous choices are at issue and when the circumstances are blurred and confused both epistemically and motivationally. Setting apart simple miscalculation from genuine uncertainty, and dishonesty

³I say most and not all cases, for in the literature two different forms of SD have been distinguished: straight SD and twisted SD. The first type refers to cases when the self-deceptive belief corresponds to the wish of the subject. The second type refers to cases when the self-deceptive belief runs contrary to the wish of the subject, as the jealous husband who falsely become convinced of his wife's infidelity without any supportive evidence (Galeotti 2016).

and duplicity from SD is often a difficult task, as I have experienced in the analysis of cases, for all these elements are usually present and entangled. Yet, the analytical category of SD helps to dispel the fogs giving each element its due, and, without supposing that a single decision can change a long-lasting trend in foreign policy, it helps assessing the consequences and responsibility of that specific decision.

A few comments (Moore, Baderin, Erez) have raised the issue of evidence, though in different directions: some affirming that evidence is always value-laden and perspectival, some pointing out lack of sufficient evidence for my empirical claim concerning the circumstances of SD, and some stressing that politicians are not researchers assessing data, but contribute to mould political reality with their actions. I have already responded to each of these remarks above; now I want to take up the role of evidence in SD and its related problems. Clearly, we can detect SD only with reference to the evidence that the agent has available, and which points in the opposite direction concerning the self-deceptive belief. A belief is self-deceptive only if it runs contrary to the evidence the agent has in front of her eyes. For instance, Anna notices that her husband is spending more and more evenings in the office, is going to conference in the weekends, closes himself in his study to make telephone calls and receives texts all night. Disquieted by her husband's behaviour, Anna ruminates over that evidence, focusing on a single positive example of her husband's caring attitude, dismissing or explaining away the problematic ones, finally ending up with the belief that her marriage is just fine and her husband caring and present. Friends and relatives, to whom Ann confided her husband's suspicious habits, conclude that she is self-deceived. In other words, there cannot be SD, if there is no evidence, available to the subject and pointing against the belief the subject has formed or has retained. This evidence needs not be conclusive, as in Anna's case, yet it must relevantly weight against the belief held by the subject, belief that appears to be induced by the wish biasing the data rather than by a correct processing of data. Hence, the notion of contrary evidence, available at the subject, is crucial for SD to take place as well as for SD to be detected by observers. This notion, however, does not commit me either to a specific view of evidence, for example is compatible with the idea of a perspectival approach, or to a specific view of the observer as impartial in absolute terms. The observers simply do not share the agent's wish, for which the available evidence represents a threat. I stress the adjective 'available', for what counts for SD is what the subject has in front of her, nor the objective evidence from an abstract viewpoint.

Finally, I want to remark that the second part of the book focused on SD in politics and on exemplary cases, is not independent of the first, concerning the philosophical discussion. Many political analysts have already referred to SD in order to explain policies that were both misguided and deceptive, yet their reference to a commonsense notion of SD has not had a relevant impact in political analysis. I contend that a well-specified concept of SD is in fact necessary for its fruitful application in politics, for only a rigorous conception of SD makes it possible to draw distinctions in the web of unwarranted beliefs, unexamined assumption, stubborn convictions and to pick out the actual episodes of SD. The analysis of the philosophical literature has led me to conclude that no single set of necessary and sufficient conditions for SD could be enlisted, given the wide variety of SD phenomenology and its fuzzy boundary. If SD is to apply to social and political reality, the concept needs to be tailored to its real complexity and balance comprehensiveness with rigour. In such a spirit, I have proposed considering three minimal necessary conditions which any kind of

SD must meet, and which jointly enable the observer to identify SD cases with a high degree of probability. They are: (1) the contextual condition, comprising the circumstances conducive to SD; (2) the cognitive condition, picking the flaws in the reasoning and the biases in the data treatment leading to the false conclusion; (3) the motivational condition, referring to the desires, aims and emotions driving the agent, in the absence of which the flows could reasonably be avoided. As I said before, the contextual condition always concerns some evidence, available to the agent, which is threatening certain crucial wish, conviction, or aim of the agents, so that the wish, conviction, aim becomes emotionally overloaded. How such a condition can be instantiated, it changes from personal to political cases and from one type to another, but the elements of emotional pressure and evidence contrary to the wish or aim of the agent are always present.

Without claiming to have solved the political problem of deception, on the one hand, and of misperception, on the other, I hope to have contributed to the understanding of a phenomenon where misperception turns into deception of the public. I have also suggested a possible line of intervention along the line of pre-commitment (Elster 1979), which however requires further reflection and a more precise proposal. Yet, before working out prophylactic measures for the SD of decision makers in politics, it is crucial that the problem must first be acknowledged.

Acknowledgments

I thank Alasia Nuti and Gabriele Badano for organizing the workshop on *Political Self-Deception* in York University. I also thank all the participants, Neil Manson, Alfred Moore, Alice Baderin, Lion Erez, Shaul Miterpunkt and Federica Liveriero for their thoughtful comments.

Disclosure statement

No potential conflict of interest was reported by the author.

ORCID

Anna Elisabetta Galeotti  <http://orcid.org/0000-0002-6142-5622>

References

- Bond, G., S. Schewe, A. Snyder, and L. Speller. 2019. "Reality Monitoring in Politics." In *The Palgrave Handbook of Deceptive Communication*, edited by T. Docan-Morgan, 953–968. Cham: Palgrave MacMillan.
- Cassam, Q. 2019. *Vices of the Mind: From the Intellectual to the Political*. Oxford: Oxford University Press.
- Elster, J. 1979. *Ulysses and the Sirens*. Cambridge: Cambridge University Press.
- Elster, J. 1983. *Sour Grapes*. Cambridge: Cambridge University Press.
- Fairweather, A., and L. Zagzebski, eds. 2001. *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*. Oxford: Oxford University Press.
- Galeotti, A. E. 2016. "Straight and Twisted Self-Deception." *Phenomenology and Mind* 11: 90–99.
- Galeotti, A. E. 2018. *Political Self-Deception*. Cambridge: Cambridge University Press.
- Jacobsen, C., T. Fosgaard, and D. Pascual-Ezama. 2018. "Why Do We Lie? A Practical Guide to the Dishonesty Literature." *Journal of Economic Surveys* 32: 357–387. doi:10.1111/joes.12204.

- Janis, I. 1982. *Groupthink: Psychological Study of Policy Decisions and Fiascos*. Boston: Houghton Mifflin.
- Landemore, H. 2012. *Democratic Reason. Collective Intelligence and Democratic Rule*. Princeton: Princeton University Press.
- MacNamara, R. 1995. *In Retrospect. The Tragedy and Lessons from the Vietnam*. New York: Random House.
- Nozick, R. 1974. *Anarchy, State and Utopia*. New York: Basic Book.
- Nozick, R. 1977. "On Austrian Methodology." *Synthese* 36: 353–392. doi:[10.1007/BF00486025](https://doi.org/10.1007/BF00486025).
- Sperber, D., and H. Mercier. 2011. "Why Do Humans Reason? Arguments for an Argumentative Theory." *Behavioral and Brain Science* 34: 57–111. doi:[10.1017/S0140525X10000968](https://doi.org/10.1017/S0140525X10000968).
- Zagzebski, L. 1996. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundation of Knowledge*. Cambridge: Cambridge University Press.