

Towards a Holistic View of Self-Deception in Kant's Moral Psychology

MARIA EUGÊNIA ZANCHET •

University of Bayreuth, Germany

Abstract

In his notable account of lying in the *Doctrine of Virtue*, Kant draws a parallel between self-deception and external lying, and argues that the agent who lies throws away her personality and dignity. Challenged by many commentators, this explanatory strategy may suggest that Kant's prohibition of deception would be motivated by a contentious teleological principle. In my account, I reject this suggestion and further show that this parallel can help us better understand the nature of self-deception. By borrowing elements from outside of Kant's treatment of self-deception in the *Doctrine of Virtue*, this paper aims to offer an account of Kant's strong condemnation of self-deception, while showing that what is at stake in cases of deception goes far beyond teleological principles. I contend that taking seriously the parallel between lying and self-deception is crucial for avoiding the trap of falling into teleological claims and that, in contrast to what some commentators suggest, the parallel is key to understanding self-deception.

Keywords

Kant's ethics; self-deception; external lying; teleology; personality; dignity.

• Universität Bayreuth, me@mariaeugeniazanchet.com

Introduction

In the section on lying in the *Doctrine of Virtue*, Kant presents what has traditionally been understood as his most systematized account of self-deception. He defines self-deception as *a form of lie* one tells oneself. This definition seems to imply a parallel and yet a *mirroring*¹ of self-deception on the external lie, since both self-deception and external lying are expressions of the same phenomenon (i.e., lying).

Yet, to mirror self-deception on external lying imposes a difficulty regarding the nature of self-deception, which is acknowledged by Kant. According to him:

"[i]t is easy to show that the human being is actually guilty of many inner lies, but it seems more difficult to explain *how they are possible*; for a lie requires a second person whom one intends to deceive, whereas to deceive oneself on purpose seems to contain a contradiction" (MS 6:430).

Nevertheless, although Kant makes explicit his awareness of the difficulties surrounding the nature of self-deception, his account of it remains controversial, as for a persistently challenging phenomenon such as self-deception,² one would expect a more explicitly systematized account for his ethics, especially given the centrality of the moral duty to know oneself (MS 2017:441) for his ethics.

But while the very lack of a unified and structured account by Kant himself leads us to wonder about the possible reasons for his laconic treatment of self-deception,³ some of the controversies have been challenged; and more recently, the overfocus on the parallel between self-deception and external lying has been criticized in the literature.

¹ In this paper I use the expressions "modeled on" and "mirrored from" to mean that self-deception inherits its functioning from external lying. This terminology is borrowed from contemporary debates on self-deception, which either endorse or challenge the strategy of modeling self-deception on intentional interpersonal deception. I preserve the idea of "parallel" between those phenomena to mean that they rest on the same grounds in a more general sense.

² Once modeled on the external lie, self-deception is a persistently challenging phenomenon as it requires that the same person simultaneously holds contradictory beliefs. This aspect of self-deception is addressed in the literature under the heading of the Static Paradox. For a comprehensive account addressing this paradox, see Mele 2001, pp. 50 – 75.

³ Additionally, Kant makes scarce use of the very concept of self-deception compared to the extent to which he discusses it from metaphors or from the description of one or more of its components, such as "rationalization" (ANTH 2007: 201; 266.), "dishonesty by which we throw dust in our own eyes" (RGV 2019: 38), frailty (MS 2017:430, RGV 2019: 38), lie to oneself, inward deceit (RVG 2019: 43).

Two approaches are worth highlighting. Papish (2018) suggests that mirroring self-deception from external lying is hardly helpful, as assessing self-deception as a category of deception in general is "ill-conceived", resulting in a view of self-deception that overlooks its epistemic aspects. According to Papish, self-deception for Kant arises once agents infringe norms of belief formation; because agents are unable to change or deny a certain cognition, they might resort to rationalization mechanisms to deflect their own attention, thus focusing on some other "minimally grounded cognition" (Papish, 2018, p. 73), which flouts the rules that are normally observed during evidence-gathering. She believes that "structural differences between deception of oneself and deception of others" (p. 71) make internal lying untenable for being a good interpretive key for self-deception.

Along the same lines, Sticker (2021) criticizes Kant's explanatory strategy by claiming that the distinction between *homo noumenon* and *homo phaenomenon* used by Kant to account for self-deception is troublesome. He suggests that in acknowledging the need for a mental partitioning promoted by a phenomenon that demands both a deceptive and a deceived self, Kant mistakenly uses the distinction between the *hominis* to account for self-deception. For Sticker, in approaching self-deception from external lying, Kant "gets it almost completely the wrong way around" (Sticker, 2021, p. 36) when he suggests that *homo noumenon* engages in self-deception by "communicating in a deceptive way with other agents," thus using *homo phaenomenon* as a mere means. Sticker then moves on to the distinction between sensuous and rational nature as a better path Kant could have taken to explain self-deception within his own framework.

Thus, although Papish and Sticker take different stances on Kant's position concerning self-deception⁴, they both maintain that overemphasizing the parallel between self-deception and external lying might be problematic.⁵

In this paper, I argue that although the framework Kant draws in the *Doctrine of Virtue* is unable to provide a sufficiently unambiguous account of self-deception, paying deep

⁴ Specifically, while Papish (2018, pp. 06, 70) argues that Kant was aware that self-deception cannot be regarded as a form of external lying, Sticker (2021, p. 26) focuses on demonstrating that Kant is mistaken to believe that the difference between *homines* is the more plausible way to explicate self-deception.

⁵ "Like me, Papish (2018: ch. 3) believes that the internal lie is not a good way to understand self-deception." Sticker, 2021, p. 24f.

attention to this parallel is central to understanding the place of self-deception in Kant's moral psychology. In this sense, with respect to Papish and Sticker's accounts, my analysis fulfills a complementary role, since it analogously concentrates on the problems with entertaining a reading that centers upon modeling self-deception on external lying. However, my account advances further theoretical aspects, as it aims to *explore* this parallel, in particular, through investigation of concepts such as personality, dignity and humanity, which I believe are the reasons Kant forged such a parallel in the first place.

In the first section, I reconstruct Kant's account of self-deception as a form of lying to oneself in the *Doctrine of Virtue*. In exploring an interpretive divergence, I hold with regard to Sticker's reading (2021, pp. 24 - 25, 36), I devise a novel account according to which Kant explicates self-deception from one's assessment over one's own epistemic attitudes. Next, I point to a possible tension arising out of the superfocus in the parallel; I discuss how modeling self-deception on external lying adds up to the premises of what I will describe as the *teleological claim about self-deception*. This claim can be inferred from textual evidence suggesting that external lying represents a violation of a natural teleological principle, that is, a principle according to which everything in nature has its own proper end, a natural purposiveness or *telos*. This points to a reading according to which, by analogy, self-deception would represent a violation of the same sort⁶.

In the second section, I look closely at this violation, which, along with Kant's strong remarks against lying and self-deception, marks the need for the discussion of the concepts Kant applies to the violation brought about by both external lying and self-deception, namely, the concepts of personality, dignity, and ultimately, humanity. My aim in this section is to explain to what extent one can attribute a teleological appeal to Kant's ethics when it comes to self-deception.

By the end of this paper, it should be clear that despite the problems that may arise from overfocusing on the parallel between self-deception and external lying, understanding the role of self-deception in Kant's ethics requires a sharp grasp of the violation involved in *both* self-deception and external lying. While a treatment of self-deception that focuses on its

⁶ The reading according to which self-deception represents a teleological violation remains unaddressed in the literature. See note 17.

mere modulation on external lying might result in a teleological misinterpretation, avoiding such misinterpretation requires precisely diving into the parallel between these two phenomena.

Section 01.

The Parallel Between External Lying and Self-Deception

The §9 of *Doctrine of Virtue* (MS 2017:429 - 431) is concerned with the ethical duty not to lie and the harms associated with the failure to carry out this duty. This discussion opens with the strong position that lying in general is the "*greatest* violation of a human being's duty to himself regarded merely as a moral being" (MS 6:429, my emphasis), which amounts to a breach of the "*humanity in [one's] own person*", a statement that is not exactly unexpected for those familiar with Kant's traditional position on lying or his classic example of the murderer at the door.

Alongside his account on lying, Kant discusses what he refers to as internal lying [*innere Lüge*], a phenomenon that falls under the scope of what we understand as self-deception [*Selbsttäuschung*]⁷. He explicates that we tell an external lie (*mendacium externum*) when we make declarations [*Erklärungen*]⁸ contrary to our beliefs directed at other persons, who in turn are led to believe the truth of those declarations; internal lies, on the other hand, amount for "insincerity in [one's] declarations, which a human being perpetuates upon [oneself]" (MS 2017:431).

Yet, in the the *Doctrine of Virtue*, more than a merely different form of lie, internal lying (henceforth self-deception) seems to be *modeled*, i.e., *mirrored* from external lying, meaning that self-deception follows the same schema as external lying. Just as external lying, self-deception seemingly involves two persons: the deceiver and the deceived, and while it amounts to an individual phenomenon, *viz.*, something that happens within one's own self,

⁷ While to assume that Kant's account regarding the duty not to lie in *the Doctrine of Virtue* exhausts his claims on self-deception is arguably mistaken, that discussion is, however, one of Kant's most systematic accounts of self-deception.

⁸ Making a statement is a condition of the possibility of lying. This excludes, for example, the idea of "lying by omission" since this modality of false statement [*Feststellung*] is necessarily non-declarative. Kant holds in his letter to Maria von Herbert (CORR 1999: 332) that only lack of sincerity is morally culpable. See also VE 1997: 62 for an account on "joking lies".

the two "persons" Kant requires to make sense of self-deception⁹ must be within the same agent. This dual psychological structure is briefly pointed out by Kant when he discusses the ethical aspects of self-deception while acknowledging its paradoxical nature:

It is easy to show that the human being is actually guilty of many inner lies, but it seems more difficult to explain *how they are possible*, for a lie requires a *second person* whom one intends to deceive, whereas *to deceive oneself on purpose* seems to contain a contradiction. (MS:430, my emphasis)

This passage indicates that Kant is mindful that once mirrored in external lying, self-deception results in a seeming contradiction. Thus, although this passage seems to suggest that providing an argument to account for the nature of self-deception is not within Kant's agenda, further textual evidence suggests otherwise.

A key point Kant raises about the nature of lying and self-deception lies on a metaphysical premise and concerns the interaction between practical standpoints. He says:

The human being as a moral being (*homo noumenon*) cannot use himself as a natural being (*homo phaenomenon*) as a mere means (a speaking machine), as if his natural being were not bound to *the inner end (of communicating thoughts)*, but is bound to the condition of using himself as a natural being in agreement with the declaration (*declaratio*) of his moral being and is under obligation to himself to *truthfulness*. (MS 2017: 430)

In this passage Kant holds that as a moral being (*homo noumenon*), the human being is not capable of [*kann (...) nicht*] using himself as a mere means¹⁰. In the context of the duty not to lie, this amounts to claiming that the human being is not capable of lying or deceiving oneself since in this respect (as *homo noumenon*) one is already bound not to lie to oneself.

These remarks throw light on self-deception because they seem to explain what is at stake from a metaphysical standpoint, meaning, by engaging in self-deception a human being uses oneself as a mere means. However, as noted by Sticker (2021), it is trivial to state that as *homo noumenon*, human beings are incapable of using themselves as natural beings¹¹, for

⁹ The relevant consequence of modeling self-deception on external or interpersonal deception is that it results in a dual-belief requirement, meaning that one must simultaneously believe P and ~P. This explanatory challenge goes back to the paradoxical character of self-deception, widely covered in contemporary literature. See for instance Mele (1983, 2001) and Van Leuween (2013). For an account that dissolves this paradox by rejecting the double-belief requirement, see especially Fernández (2013).

¹⁰ Which means treating oneself at the expense of one's own humanity, i.e., as a thing. See GMS 2011: 429.

¹¹ This passage is discussed by Sticker (2021, pp. 23 - 26), who problematizes the strength of the *homines* distinction when it comes to accounting for self-deception. In my reading, in order to assess why Kant draws this distinction, one needs to allow for the role of concepts such as personality and dignity in his account of

"the *homo noumenon* cannot do anything immoral anyway, since it is our legislating reason" (Sticker, 2021, p. 24).

However, beyond asserting a triviality, Kant is making an important point here, namely, that moral transgression rather occurs with regard to one's phenomenological aspect. This reading frames this passage as containing a normative claim. This is better elucidated once we take into account the distinction between *homines* in light of *the way* one uses one's rational capacities.

That an agent (as *homo noumenon*) uses one's legislating reason for immoral purposes is, as a matter of fact, beyond one's capacities - and thus beyond one's control. In effect, anything outside the scope of human capacities and control cannot be considered a duty (RGV 2019: 47; MS). This further reinforces that the passage should not be read as a warning on how one is required to bring into balance one's own metaphysical parts. However, there is something that is indeed within one's control, namely, *to see oneself*¹² as "subject of the moral lawgiving which proceeds from the concept of freedom and in which he is subject to a law he gives to himself," which in turn implies a certain use an agent makes of one's own practical reason. This use, Kant explains, compels one to regard oneself as well as other human beings *as ends in themselves*. As he states in the *Groundwork*¹³, "a human being (...) exists as an end in itself (...) but *must* in all its actions, whether directed towards itself or also to other rational beings, *be considered* at the same time as an end" (GMS 2011: 428f, my emphasis).

Moreover, the normative claim Kant makes in this passage can be better appreciated once we pay close attention to the terminology. Notably, Kant refers to the idea of condition [*Bedingung*] in order to establish an agreement [*Übereinstimmung*] between the different aspects of one's being. Of course, the concept of condition is not inherently normative;

lying in the *Doctrine of Virtue*. I am particularly skeptical toward Sticker's criticism, for I believe that Kant uses the *homines* distinction to draw attention to how one ought to regard *oneself* when it comes to one's duty of truthfulness.

¹² "When we *think of ourselves* as free we transfer ourselves into the world of understanding as members of it and cognize autonomy of the will along with its consequence, morality; but if we think of ourselves as put under obligation *we regard ourselves* as belonging to the world of sense and yet *at the same time to the world of the understanding*." (GMS, 2011: 443)

¹³ In the *Doctrine of Virtue*, this additionally means that in order to make the proper use of one's practical reason, one must conceive of oneself as being under someone else's will, since conceiving another's will is a condition for us to make the idea of obligation intuitive for ourselves (MS 2017: 487).

however, in this case, where the issue in play is precisely the moral outcome resulting from lying,¹⁴ Kant seems committed to the claim that bringing the *homines* into agreement is a condition that must be met if one is to make appropriate use of one's moral capacities. Thus, to put it another way, that an agent thinks of herself as a moral being, i.e., as freely legislating over the principles of her actions, is a necessary condition so that she can treat herself and others accordingly; this includes, in this context, avoiding lying or engaging in self-deception. In contrast, once she fails to see herself as a member of the intelligible world (meaning thereby, failing to see herself as *homo noumenon*) she will also fail to make actual this aspect of herself as a person¹⁵. Thus, what Kant does in this passage is to stress, from a metaphysical point of view, the existence of an agent's duty to oneself "in regard to its substance" (VE 27: 601), namely, that an agent must see, regard, or think about oneself as a moral being.

Yet, in attempting to clarify the nature of self-deception Kant provides us with a few examples of a self-deceptive agent regarding his beliefs about the existence of God:

Someone tells an inner lie, for example, if he professes belief in a future judge of the world, although he really finds no such belief within himself but persuades himself that it could do no harm and might even be useful to profess in his thoughts to one who scrutinizes hearts a belief in such a judge, in order to win his favor in case he should exist. Someone also lies if, having no doubt about the existence of this future judge, he still flatters himself that he inwardly reveres his law, though the only incentive he feels is fear of punishment. (MS 6:430)

In the first case, the agent deceives himself by believing in something which he initially does not believe, but which he thinks is to some extent harmless and worthy of endorsement. In the second, the self-deception results from the agent's misinterpretation about his own incentives.

It is worth noting that in both examples, Kant refers to epistemic strategies to account for the possibility of self-deception, that is, to explicate how self-deception can occur from the assessment over one's own epistemic attitudes. These epistemic strategies of insincerity amount to one's impurity in the declarations one makes before one's own conscience, i.e.,

¹⁴ It should be noted that it is precisely in the previous paragraph that Kant constructs lying as resulting in the renunciation of one's personality.

¹⁵ That is, as a person who embodies legislating reason, that is, humanity and dignity. These concepts will be discussed below.

before one's "inner judge", being this, the second person Kant requires to make sense of self-deception.

In *Religion*, also in discussing one's belief in the existence of God, Kant challenges the principle whereby "it's advisable to believe too much rather than too little" (RGV 2019: 188). He elaborates a harsh critique of this principle based on the same justification addressed in the *Doctrine of Virtue*. For him, to use such a "safety maxim", that is, to force oneself to believe in something out of convenience is a violation of conscience amounting to dishonesty in one's pretense.

As such, from the epistemic strategies just mentioned, both self-deception and external lying embody untruthfulness since by lying one violates, to some extent, the beliefs one professes¹⁶. Analogously, failing to regard oneself as a moral being seems to be based on an epistemic distortion, albeit not concerning discrete epistemic attitudes, but rather one's assessment regarding one's metaphysical standing. By virtue of their epistemic quality, these violations are, in this sense, opposed to truthfulness (MS 6:429), which is why both forms of lying are objects of the strictest censure (MS 6:430).

The Teleological Claim about Self-Deception

Overfocusing on the parallel between these two forms of lying may, however, lead to misconceptions of self-deception. This is among the premises that result in what I discuss under the name of the teleological claim about self-deception. The (henceforth) teleological claim emerges from Kant's treatment of external lying (MS 6:429), which is defined as "communication of one's thoughts to someone through words that yet (intentionally) contain

¹⁶ The way in which external lying and self-deception function is, admittedly, different. In the case of lying, seeing how lying violates truthfulness is unproblematic. For example, when an agent promises to pay a debt even though she has no intention of paying, she deceives the person to whom she has lied. However, for cases of self-deception the violation of one's own belief needs to be more subtle, otherwise it would not result in self-deception. It is precisely this dual belief requirement that creates the apparent contradiction Kant mentions. However, there are strategies to avoid the dual belief requirement presumably involved in self-deception. Instances of such strategies are attributing epistemic flexibility, by means of postulating different levels of belief, such as *deep*, *stated* and *experienced* belief (Mijović-Prelec & Prelec, 2010); predicting non-doxastic attitudes as, for example, *S suspects* that *P*; or arguing that what is at stake is a shift of focus. On the latter argument, of which mine concurs, see Papish (2018, ch. 03). Yet these are not the only ways to obviate this seemingly inevitable contradiction. The very *Doctrine of the Elements of Ethics* is anchored in Kant's answer to a problem that emerges when one tries to conceive how "duties to myself" are possible (see MS 2017: 418). His solution to the problem of being passively constrained and actively constrained is precisely the postulation of two aspects of the same agent, *homo noumenon* and *homo phaenomenon*.

the contrary of what the speaker thinks on the subject” and is qualified as violating “the natural purposiveness of the speaker's capacity to communicate his thoughts.”¹⁷

In this passage, Kant seems to claim that lying violates the purpose in communicating one's thoughts to another. This violation results from the conflict between two ends: the end of lying and the natural end (natural purposiveness) of communication. It also reflects a limitation or misuse of one's capacities. The terminology employed by Kant in this passage reinforces the appeal to a teleological interpretation, since by putting together the concepts of end, purposiveness, and nature, Kant seems to intend a reference to the idea of *telos* or final end.

Kant's supposedly naturalistic attitude emphasized by such terminology is especially compelling for those familiar with his considerations on the nature of the will in the *Groundwork*, according to which, if nature has endowed us with practical reason, it follows that practical reason must have a final end. These considerations encompass the idea that, “[i]n the natural predispositions of an organized being, i.e., one arranged purposively for life, we assume as a principle that no organ will be found in it for any end that is not also the most fitting for it and the most suitable” (GMS 2011:395).

Furthermore, if one takes a closer inspection on Kant's structure of duties, one can see that there is a parallel between external lying and the violations related with the other duties, on which the teleological reference is even stronger. According to Kant,

Just as love of life is destined by nature to preserve the person, so sexual love is destined by it to preserve the species; in other words, each of these is a *natural end*, by which is understood that connection of a cause with an effect in which, although no understanding is ascribed to the cause, it is still thought by analogy with an intelligent cause, and so as if is produced human beings on purpose. (MS 6:424)

Violations such as suicide or nonprocreative sex seem to violate the same teleological principle, since the duties opposed to these violations correspond to alleged natural ends.

¹⁷ Most explicitly, Timmermann (2000, p. 280) points to the possibility of a teleological reading of Kant's account of lying. Gregor (1963, p. 139) and Denis (2012, pp. 104 - 110) challenge Kant's appeal to teleological principles in his taxonomy of duties. Dietz (2002) in turn draws a positive relationship between lying and the violation of teleological principles. In my understanding, this is due to the overly strong emphasis she places on lying mostly as a wrongful or misuse of language. In missing the point, Dietz is led to claim that Kant holds a conception of language which admittedly has a “single function, that of true communication” (p. 99), a claim that is sound only if one assumes teleological premises.

Therefore, there seems to be a parallel between external lying and the other duties Kant discusses in the *Doctrine of Virtue*, as the violations that arise from them seem to be equally bound to a teleological principle, since they violate natural ends.

Along these lines, if self-deception mirrors external lying, then analogously, we are left to situate the violation that results from self-deception within the teleological field. This in turn makes room for the idea that self-deception is wrong in virtue of violating the natural purposiveness of truthfulness a human being has towards oneself.

Thus, the teleological claim can be described as follows: Self-deception is a form of lying in the sense that it mirrors and therefore follows the same structure as external lying; external lying, like nonprocreative sex or suicide, seems to involve a violation of the proper use of human capacities, that is, a teleological violation. Therefore, self-deception represents a violation of the same nature.

While not elaborating or directly discussing the problem involving the teleological claim of self-deception, some commentators provide us with helpful elements for solving it. For example, while arguing that there are four senses in which Kant's moral theory can indeed be considered teleological, Guyer (2002) shows that the principle outlining the existence of a proper use of human faculties does not offer in Kant any normative function and therefore Kant's treatment of cases such as suicide and nonprocreative sex must be considered merely heuristic rather than properly teleological.¹⁸ The latter premise that results in the teleological claim about self-deception (namely, that external lying entails a teleological violation) may be tackled from Guyer's remarks, which allow us to argue that while Kant's moral theory may be considered teleological "virtually from the outside," the discrete cases of such violations, insofar as they are based on "the teleological assumption that everything in nature has a purpose" (Guyer, 2002, p. 182) have a merely heuristic role.

The first premise (that self-deception mirrors external lying) can in turn be countered by arguments from the commentators already mentioned, who emphasize the problems in construing self-deception as a form of external lying.

¹⁸ This is because assuming that "it is immoral to adopt an end other than that nature intends for us (...) has no justification" insofar as it proves to be "incompatible with [Kant's] fundamental principle of unconditional value of human freedom" (Guyer, 2002, pp. 180 - 181)

However, there are similarities between the two phenomena that are vividly described by Kant in his treatment of lying in the *Doctrine of Virtue*. The strongest link between the two lies in the violation they pose. Both forms of lying seem to be regarded by Kant as leading to the relinquishment of one's personality [1] and the annihilation of the dignity [2] of the agent who now "has even less worth than if he were a mere thing." (MS 2017: 249)

The firm opposition with which Kant stands against lying is as puzzling as its consequence: whoever lies loses their dignity and personality. Yet it is precisely this attitude that typically strikes us as odd. After all, how is it possible to cast doubt on a characteristic from which we can derive our worth as persons merely as a result of an ordinary and pervasive behavior such as lying?

It is specifically in trying to answer this question in light of the considerations I have just raised¹⁹ that the teleological claim appears attractive; postulating a natural teleological principle would greatly contribute to explain these violations' severity.

Following Papish and Sticker's claims, I also assume that interpreting Kant's mirroring of self-deception from external lying might be problematic. However, I believe that Kant has a point in tracing such a parallel, and the centrality of the consequences Kant assigns to the violations entailed by *both* phenomena support this way to frame his account. In the next lines, I thus explore the parallel between these two forms of lying. I aim to devise a strategy that avoids a reading of the *Doctrine of Virtue* overly focused on this relationship that, by extension, takes part in the teleological claim. To explain what Kant might have in mind in that framework, I seek resources beyond the *Doctrine of Virtue* discussion of lying that allow us to make sense of what is ultimately at stake when it comes to lying and self-deception. I take the endeavor to explain the forcefulness in which Kant stands against lying and self-deception as a guideline and look closely at the concepts that are at play when Kant discusses these phenomena, namely, (1) personality and (2) dignity. Discussing them at length will be critical towards a better understanding of the place of self-deception in Kant's moral psychology.

¹⁹ These are: Kant's position that lying amounts to a violation of the purpose of communication; his claims in the *Doctrine of Virtue* about the consequences entailed by lying and self-deception; and the terminology he employs to refer to these violations. In addition, the teleological elements outside of his account of lying, which, as discussed, are more explicit in his discussions on nonprocreative sex and suicide.

Section 02.
Personality, Dignity, and Self-Deception

Kant is adamant when elaborating on the consequences of violating one's duties to oneself and, in particular, of lying. In his treatment of lying (MS 2017: 429) Kant refers to the renunciation [*Verzichttuung*] of one's personality that accompanies the annihilation [*Vernichtung*] of one's dignity as a human being. He also mentions the idea of annihilating [*zernichten*] the subject of morality in one's own person (MS 2017: 423). Further on, he states that by using oneself "merely as a means to satisfy [one's] animal impulse" (MS 2017: 425) one surrenders [*aufgeben*] or throws away [*wegwerfen*] one's personality. Furthermore, that false humility amounts to the degradation [*Abwürdigung*] of one's personality (MS 2017: 436); in addition, in the context of what it means to be a useful member of the world, Kant states that it encompasses a duty to not degrade [*abwürdigem*] humanity in one's own person (MS 2017: 446).

But although he refers to the loss of personality as one of the consequences of violating the duty to oneself not to lie, Kant elaborates the classical definition of personality elsewhere. In the *Critique of Practical Reason*, Kant states that personality means the "freedom and independence from the mechanism of all nature yet regarded at the same time as a power of a being subject to pure practical laws that are peculiar to it." (KpV 2002: 87) Personality, or independence from sensible impulses, is bound up with a predisposition in us bearing the same name. This and two other predispositions to the good in human nature²⁰ are listed and elaborated by Kant in the first chapter of *Religion* (cf. RGV 2019: 26 – 28). For Kant, the predisposition to personality has its basis in practical reason, thus delimiting our end as rational human beings insofar as we act morally. Unlike the other predispositions which have vices associated with them, the predisposition to personality allows respect for the moral law to stand as a sufficient incentive for the power of choice²¹. Personality amounts to "the idea

²⁰ Which are the predispositions to animality and humanity (RGV 2019: 28).

²¹ I follow the remarks of Pasternack (2013, p. 96), who emphasizes that in contrast to the other two predispositions, "the Predisposition to Personality is without a dark side. It can, of course, be ignored, but it cannot be corrupted." This predisposition is rather connected with the "germ of goodness" (RGV 2019: 45), which remains always pure.

of the moral law alone, together with the respect that is inseparable from it.” (RGV 2019: 28)

While it is not possible to get rid of our predisposition to personality,²² since it determines our nature as moral beings,²³ Kant suggests in the *Doctrine of Virtue* that the same does not apply when it comes to personality itself, for it can be renounced, annihilated, thrown away, or degraded. The following lines are concerned with shedding light on this possibility.

Renouncing One's Personality and Violating One's Dignity

To be a person is to be a living being endowed with a moral personality.²⁴ Unlike a thing, a person is a subject “whose actions can be imputed to him.” (MS 2017: 223) This ability for action, - and the possibility of moral accountability, goes back to the duplicity of our nature, which is sensible but at the same time intelligible, whose will needs to be constrained by the law so that good deeds can result from it. According to Kant:

We conceive of man first of all as an ideal, as he ought to be and can be, merely according to reason, and call this Idea *homo noumenon*; this being is thought of in relation to another, as though the latter were restrained by him; this is man in the state of sensibility, who is called *homo phenomenon*. (VE, 1997: 593)

As *homo noumenon*, we are only a “personified *idea*”, namely, the idea of a subject under the moral law, whereas, as embodied persons, we are “affected by the feelings of pleasure and pain” (VE, 1997: 593)²⁵. Both of these aspects belong together to the idea of personality, understood as “freedom of a rational being *under moral laws*” (MS 2017: 223, my emphasis), because acting morally depends on the necessitation of our will by the moral law.

The morality (implied in the idea of personality) is, “the condition under which alone a rational being can be an end in itself; because it is possible only by this to be a legislating member in the kingdom of ends”. (GMS 2014: 435) Kant concludes that the dignity of humanity lies in the capacity for morality. To put it another way, personality, as the characteristic of rational beings inasmuch as they are capable of being affected by and

²² “Freilich muß hierbei *vorausgesetzt werden*, daß ein Keim des Guten in seiner ganzen Reinigkeit übriggeblieben, nicht vertilgt oder verderbt werden konnte” (RGV 2011: 45, my emphasis).

²³ It is referred to, together with the two others, as “*original*” predisposition. Cf. RGV 2019: 28.

²⁴ In addition to psychological personality, which traces back to the “ability to be conscious of one's identity in different conditions of one's existence” (MS 2017: 223).

²⁵ Cf. also KvP 2002: 86 – 87.

adopting the moral law as their main incentive, gives human beings their worth or dignity [*Würde*]. In sum, the moral capacity of rational beings is the basis of a person's dignity²⁶.

Yet, renouncing one's own personality is not something that occurs in isolation. When a person tells a lie, in addition to renouncing her personality she also violates her dignity as a human person, for the root of her dignity lies in the ability to provide for herself moral principles, i.e., in her personality. That is, while a person who lies "has even less worth as if he were a mere thing" (MS 2017: 429) he also "violates the dignity of humanity of his own person," (MS 2017: 429) degrading himself "far below the animals." (ANTH 2007: 489)

However, concluding that the renunciation of dignity and personality take place simultaneously is still not the same as demonstrating that such renunciations are possible. If they are connected to our predispositions and thus to human nature, that a lying person gives up her dignity and personality still seems to contradict some common intuitions concerning our constitutive features.²⁷ In other words, if dignity and personality are intrinsic features of human beings as we typically hold, it becomes difficult to see how one can abdicate these properties. What makes it possible for an agent to acquire the "mere appearance of a human being, not a human being himself" (MS 2017: 429) remains therefore unclear.

It seems that if we are to make sense of this renunciation, then dignity and therefore personality must be something other than descriptive concepts, outlining qualities human beings do or do not *inherently* carry.

Oliver Sensen (2009) analyses Kant's use of the term dignity in different contexts. Contrary to what has been argued by other Kant scholars, Sensen contends that dignity corresponds to a relational property, notably, a property that belongs to something in relation to something else. In the case of human dignity, it can be assumed that by virtue of certain capacities, human beings possess a prerogative or elevation over other beings whose will is

²⁶ This connection is pointed out by Wood, A. (1999), who provides us with what Bayefsky (2013) calls a moral capacity argument in regard to the grounds of moral dignity.

²⁷ See for instance MS 4:463, where Kant suggests that although an agent's deeds go against duty, one cannot withdraw this person's dignity: "I cannot withdraw at least the respect that belongs to [a vicious man] in his quality as a human being, even though by his deeds he makes himself unworthy of it".

merely sensitively determined.²⁸ From this follows the idea that there are, in Sensen's words, "two stages" of the elevation of dignity. On his account, as human beings, we have an initial dignity that can be enacted to the extent that we make appropriate use of our moral capacities, but which can also, for that very reason, be violated. Herein lies the relationship between dignity and personality, to which I alluded earlier. The second stage would therefore be this actualization. When we do not bring about our dignity, that is, when we refuse to act according to moral principles and do not make appropriate use of our freedom, we fail to elevate our moral capacity. To the extent to which she "deprives [her]self of the prerogative of a moral being" (MS 2017: 420), the person who tells a lie or engages in self-deception violates the duty against herself along with the dignity of humanity in her personality.

The account drawn and advocated by Sensen makes it clear that the predisposition to personality is woven together with moral accountability. This predisposition functions as a subjective condition for the moral law to be apprehended, and therefore, as a condition of the consciousness of our freedom - namely, of the freedom of our will, or yet, of the "independence of our power of choice from determination by all incentives." (RGV 2019: 26f) Consequently, when we tell (internal or external) lies, we cease to make effective a central aspect of our rationality. In failing to regard ourselves as moral beings (*homo noumenon*), we also give up on that which engenders our moral responsibility, that is, the freedom of our power of choice.

Another aspect that makes clear the moral and epistemic strength of the violation posed by lying and self-deception is cast by Kant in the opening sentence of his account of lying in the *Doctrine of Virtue*. Telling a lie or engaging in self-deception amounts to a violation which is "contrary to *truthfulness*" (MS 2017: 429, my emphasis).

The idea of truthfulness [*Wahrhaftigkeit*] underlies Kant's entire account of these phenomena. According to Kant, truthfulness in general (*rectitude*) encompasses two further attitudes: honesty [*Ehrlichkeit*], which is truthfulness in the statements we make; and sincerity [*Redlichkeit*] when the statements we make are promises.

²⁸ Sensen delves, among other passages, into Kant's discussion of servility in the MS (2017: 434). In this discussion, what has dignity is "the moral aspect of oneself (...)" which is "elevated over the merely natural aspect of oneself" (Sensen, 2009, p. 329).

Such statements may embody truthfulness even though they are not themselves true²⁹. For example, an agent who tells a lie and yet considers herself to have behaved morally was typically not diligent enough in her self-reflection upon the statement on her own action - that is, whether or not that statement embodies truthfulness. It is precisely in this context that our inner judge comes into play. It is our inner judge, i.e., our consciousness that assesses the statements we make, whether or not they embody truthfulness (rather than whether or not they are true). Truthfulness is therefore essential to the way we assess the statements we make about the state of affairs in the world but also, and most importantly, *to the way we see or regard our moral selves*.

Therefore, in making false declarations that additionally incorporate deception, we violate a duty which is intimately connected to our self-cognition, thus compromising our ability to judge ourselves responsible for our actions. An agent who deceives oneself and rationalizes away her responsibility for her immoral deeds, evades the accountability and moral obligation that are in turn directly engendered by the fact that she is a person, i.e., that which allows one to regard oneself as *homo noumenon*. It becomes thereby clear why truthfulness, i.e., the exact opposite of lying³⁰ also indicates an obligation one has to oneself *as a moral being*.

There are, however, differences in terms of priority of self-deception over external lying. While external lies *may* also harm others, internal untruthful declarations *always* violate our self-respect (MS 2017: 404), for they aim to deceive our inner judge, both resulting in the lie *per se* but also in the awareness of that lie we have told to ourselves. In this sense, the respect a human being has for oneself is firmly grounded on truthfulness³¹. As a result of this violation, the access we have to ourselves as moral beings other than mere “speaking

²⁹ The concept of truthfulness [*Wahrhaftigkeit*] is derived from the terms ἀληθής/ἀλήθεια, which convey a disposition of character indicating aversion to deceptive behavior, such as lying or self-deception. When predicated to the agent, *truthfulness*, or in this case *the truthful agent*, has been defined as "verum dicit et veritatis adsertor est", meaning "who says the true" (Szaif, J., & Thurnherr, U., 2004, p. 42). Such etymological definition entails that the truthful agent must always have access to the modal content of her statements, thus knowing whether its content is true or false. This implication is philosophically troublesome, especially from an ethical perspective, and Kant seemed to be aware of this, as his use of the term does not imply the agent's access to the truth or falsity of her statements, but relies instead on that agent's maxim.

³⁰ "Between truthfulness and lying (which are contradictorie oppositis) there is no mean." (Cf. MS 2017: 434). On this issue, see Pinheiro Walla (2013, pp. 312 - 314).

³¹ Cf. KpV 2002: 93. Moreover, the very process of maxims-assessment by practical reason also depends on truthfulness, as held by Kant in KpV 2002: 44.

machines” (MS 2017: 430)³² is thus jeopardized, from which it follows that lying additionally entails the violation of our personality.

Self-deception is, therefore, equivalent to a renunciation that is intimately intertwined with the proper use of our mental faculties. These faculties are in turn the distinguishing feature between persons and mere things. The renunciation of our personality, therefore, goes back to the renunciation of a unique characteristic of human beings: the *moral aspect* of our dignity. Therefore, when we make declarations that incorporate deception we compromise our function as human beings, thus violating, “the highest principle of truthfulness” (MS 2017: 431), that on which one depends in order to be able to regard oneself a moral being.

This is where the meaning of Kant's statements might become misunderstood. As a matter of fact, modeling self-deception on external lying renders this renunciation even clearer. Take for instance the passage where Kant constructs a parallel between violating the perception of humanity and violating the internal end of communication discussed earlier. In this passage, Kant is adamant that the intention to deceive corrupts the proper function, i.e., *natural purposiveness* of communication between two persons. That self-deception violates self-respect is a result of the same violation, albeit applied to the agent oneself. Thus, once one assumes that Kant models self-deception on external lying, it seems reasonable to infer that the reason for Kant's depreciation of deception is that lying violates the proper function (*telos*) of communicating our thoughts and, alongside the external lie that corrupts the proper function of communicating our thoughts, self-deception corrupts truthfulness³³. Because every lie implies an initial self-deception, both phenomena are deeply problematic for Kant. Thus, as a result of violating our freedom, we become “a plaything of the mere inclinations and hence a thing.” (MS 2017: 420)

This conclusion sounds correct, and Kant's discussion of truthfulness seems to stress it. The problem lies, however, in holding that the severity of this violation arises from the idea that by lying one violates a teleological principle. In other words, what is problematic here is to assume that Kant condemns self-deception merely because, just as lying violates the end

³² This and other metaphors seem to be used by Kant to emphasize that we have remarkably little left when we violate our personality. We become mere speaking machines inasmuch as speaking machines, or mere things, are objects whose will is ultimately determined by laws of nature.

³³ Cf. MS 2017: 429 – 430.

communication to others, self-deception violates the end of communication to oneself, i.e., truthfulness.

However, as I hope to have shown in my discussion of the concepts of personality and dignity, when one engages in deception, what is at stake is something much greater than the violation of the proper function of telling the truth, taken in an essentially natural way. What is at stake is instead the violation of our capacity for morality itself, which qualifies a human being as a person, as opposed to a mere speaking machine. As Kant says, "to annihilate the subject of morality in one's own person is to *root out the existence of morality itself from the world*" (MS 2017: 423, my emphasis).

Some remarks on the interplay between humanity and personality might help us better understand the violation of morality at hand here. The concept of humanity is a key aspect in Kant's ethics, which is why it features in the second formulation of the categorical imperative³⁴ presented by Kant in the *Groundwork*. Kant's definition of humanity, that is, as an "objective end" that must be treated as an end rather than a mere means, is consistent with Kant's use of this concept throughout the *Metaphysics of Morals*.

In the *Doctrine of Right*, Kant defines humanity as grounding innate freedom, which in turn is prerogative of any human being (MS 2017:238). Notably, humanity is a property of one's capacity for freedom, and should therefore be understood as one's "personality independent of physical attributes" (MS 2017: 239).

In the *Doctrine of Virtue* the overlap between humanity and personality (through dignity) is prominent. Not only does Kant discuss personality and dignity simultaneously (as seen in his account of lying), but he also resorts to these concepts by stating that "[h]umanity itself is a dignity." He explains:

for a man cannot be used merely as a means by any man (either by others or even by himself) but must always be used at the same time as an end. It is just in this that his dignity (personality) consists, by which he raises himself above all other beings in the world that are not men and yet can be used, and so over all things. (MS 2017: 462)

³⁴ "Act that you use humanity, in your own person as well as in the person of any other, always at the same time as an end, never merely as a means" (GMS 2011: 429).

The same relationship between humanity and personality can be also appreciated in *Religion*. As discussed, Kant claims that we hold a "susceptibility to respect for the moral law" (RGV 2019: 27) by virtue of our predisposition to personality. This suggests this predisposition has two aspects: the first, "the *subjective* ground of our incorporating this respect into our maxims," (RGV 2019: 27)³⁵; and the second, "the idea of moral law alone," that is, the objective aspect of that very predisposition. This latter aspect accounts for what we may properly call "personality," which, Kant explains, "is (...) the idea of humanity considered wholly intellectually" (RGV 2019: 28).

Thus, to regard oneself or others as *homo noumenon* amounts, in that sense, to regard oneself or others in terms of their humanity (MS 2017: 295), which in turn is equivalent to regard oneself (or others) "as a person, that is, as the subject of a morally practical reason" (MS 2017: 435). Therefore, what Kant calls a person in a moral sense (*homo noumenon*), which by virtue of being an end in itself, "is exalted above any price," is analogous to the idea of humanity considered intellectually.

Accordingly, to conceive of oneself objectively, that is, to look at the objective aspect of one's own person, is a condition for treating oneself and others as ends in themselves³⁶. By self-deceiving and violating one's own personality, one additionally violates one's humanity. In the opening sentence of the discussion of lying in the *Doctrine of Virtue* Kant makes this point clear by referring to lying as "the greatest violation [against] the humanity in [one's own] person" (MS 2017:429).

Of course, from the foregoing, one might argue that the insertion of the concept of humanity as an end in itself, and its equivalence with personality, weaves a fundamentally teleological sense into Kant's injunction against self-deception. As a matter of fact, some interpreters³⁷ are sympathetic when it comes to drawing out intersections between ethics and teleology in Kant. However, being careful on this point is crucial, lest one incurs a purely natural teleology, which makes a direct appeal to a naturalistic ethics where the violation of duties is explained through the violation of teleological principles.

³⁵ That is, the fact that this predisposition points to a natural aspect of our constitution as beings whose sensibility is a condition for apprehending the moral law.

³⁶ See especially MS 2017: 379f.

³⁷ As discussed by Guyer (2002). See also Boxill (2017).

The sense in which the *Doctrine of Virtue* can actually be considered teleological is instead a strictly moral sense, within a framework used by Kant to point to the moral nature of agents as end in themselves. Their end is *moral* self-preservation, stressed in particular in Kant's discussions of the duties to oneself as a *merely moral being*, of which truthfulness is a crucial part.

Additionally, for discrete cases of self-deception it is simply wrong to claim that the violation of truthfulness lies in a violation of a teleological principle. Even duties to oneself as animals, i.e., those that refer to essentially animal impulses, do not depend exclusively on a purely naturalistic appeal. This is because even as animals, that is, as finite and natural beings, we are also endowed with two other predispositions, namely humanity and personality (RGV 2019: 26). For this reason, the distinction Kant sets forth in the first book of the *Doctrine of Virtue* is between one's duties to oneself as an animal being [*als einem animalischen Wesen*] and one's duties to oneself merely as a moral being [*bloß als einem moralischen Wesen*]. This distinction underlines the fact that when discussing the duties one has to oneself as a moral being, Kant is isolating this property, that is, letting animality out. Thus, whereas even in addressing the duties to oneself *as an animal being*, Kant uses the naturalistic principle "without harm"³⁸ only as methodological support for his claims concerning these duties, and therefore it is even less likely that violations of duties to oneself *merely as a moral being* rely (in an essentially naturalistic way) on teleological principles.

This suggests that Kant's argument does not rely upon teleological claims of any kind, which instead are merely meant to make explicit the seriousness of the violations of the duties to oneself. Kant is rather concerned that violations of such duties result in the loss of that which is a condition for all moral action, namely, the loss of personality, i.e., the (purely intellectually considered) humanity of the agent. He is therefore concerned with violations, whether of formal or material duties, that result in the agent being prevented from properly using her capacities with respect to the exercise of morality. Even more specifically, both phenomena, external lying and self-deception, preclude the possibility of setting maxims that embody truthfulness, thereby jeopardizing the very chance of deeds out of duty, a central element in Kant's ethics, which stands for the exercise of morality.

³⁸ See KU 2000: 379.

Conclusions

The growing interest of many Kant scholars in self-deception has recently placed this phenomenon at the core of Kant's ethical thought. Yet with regard to explaining Kant's views on self-deception, things may become obscure, for the parallel he draws between self-deception and external lying, traditionally regarded as his most systematized attempt to explicate self-deception, is not without difficulties.

As a matter of fact, such a parallel is involved in a number of explanatory quandaries. It may lead to problems concerning the very nature of self-deception, but it can also importantly contribute to misreadings of some of Kant's ambiguous assertions.

Out of the latter case might emerge the teleological claim I have addressed here. This would be the claim that the prohibition of lying and of self-deception is based on a teleological principle, meaning that self-deception would violate the natural end of communication, and therefore infringe the truthfulness that one must have towards oneself in order to formulate maxims that would lead to actions out of duty. The parallel between these two forms of lying would, according to this interpretation, reinforce the teleological claim.

However, Kant does not rely on this claim to establish what is wrong with self-deception. The passages in which he expresses his firm rejection of lying provide textual evidence that the violation of the proper function of telling the truth is not what is at issue in his account in the *Doctrine of Virtue*. For Kant, the violation associated with self-deception is due to the fact that such a phenomenon hinders the use of moral abilities. More specifically, self-deception impedes the agents' capacity to see themselves as moral beings, meaning to bring into effect what characterizes them as persons, rather than as things, i.e., their personality. The teleological claim, so I argue, therefore takes on a purely heuristic role.

Interpreting self-deception alongside external lying, that is, adopting an explanatory strategy that, rather than being dismissive, brings attention to this parallel, is precisely what gets us to understand the crucial point Kant lays out in the *Doctrine of Virtue* when he renders both phenomena so inextricably connected. His aim is to stress that both external lying and self-deception impose an important risk to the exercise of one's rational capacities insofar as both lies affect how one regards oneself and how one sees one's own relation to pure practical reason.

Once we realize that there is a point therein, we are invited not only to acknowledge the limitations of the framework of self-deception Kant devises in the *Doctrine of Virtue*, but most importantly, to strive for connections between aspects that Kant himself discusses outside of that framework.

More than drawing attention to the similarities between self-deception and external lying, this paper contributes to the debate on Kant's moral psychology by addressing both phenomena on the basis of the violation *both* represent. Furthermore, I have systematized and extended the application of the teleological claim to the case of self-deception, arguing for the complementarity of multiple passages when it comes to making sense of Kant's claims about self-deception in the *Doctrine of Virtue*, thus adding a new layer to the arguments for the centrality of self-deception in Kant's ethics.

Bibliographical References

Allison, H. E. (1995), "Reflections on the banality of (radical) evil: A Kantian analysis". *Graduate Faculty philosophy journal*, no. 18(2), pp. 141–158.

Allison, H. E. (2010), *Kant's theory of freedom*, Cambridge University Press, Cambridge.

Bayefsky, R. (2013), "Dignity, honour, and human rights: Kant's perspective", *Political Theory*, no. 41(6), pp. 809 – 837.

Boxill, B. (2017), Kantian Racism and Kantian Teleology, In Zach, N., *The Oxford Handbook of Philosophy and Race*, Oxford University Press, UK, pp. 44 - 53.

Dietz, S. (2002), Immanuel Kants Begründungen des Lügenverbots, in Leonhardt, R. & Rösel, (eds.), *Dürfen wir lügen*, Vandenhoeck & Ruprecht, Germany, pp. 91–115.

Fernández, J. (2013), "Self-deception and self-knowledge", *Philosophical Studies*, no. 162(2), pp. 379 –400.

Hill, T. E. (1971), "Kant on imperfect duty and supererogation", *Kant-Studien*, no. 62(1), pp. 55 – 76.

Formosa, P. (2009), "Kant on the limits of human evil", *Journal of Philosophical Research*, no. 34, pp. 189 – 214.

Gregor, M. J. (1963), *Laws of Freedom: a Study of Kant's Method of Applying the Categorical Imperative in the Metaphysik der Sitten*, Blackwell, UK.

Grenberg, J. (2010), "What Is the Enemy of Virtue?", in Lara Denis (ed.), *Kant's Metaphysics of Morals: A Critical Guide*, Cambridge University Press, UK.

Guyer, P. (2002), "Ends of reason and ends of nature: The place of teleology in Kant's ethics", *Journal of Value Inquiry*, no. 36(2-3), pp. 161 – 186.

Kant, I. (1786), *Conjectural beginning of human history* in (2007), Zöllner, G., & Louden, R. B. (eds.), *Anthropology, history, and education*, Cambridge University Press, UK.

Kant, I. (1996), *On a Supposed Right to Lie from Philanthropy*, in Gregor, M. J. (ed.), *Kant, Practical Philosophy*, Cambridge University Press, UK.

Kant, I. (1997), *Lectures on Ethics*, ed. by Heath, P. & Scheenwind, J.B., Cambridge University Press, UK.

Kant, I. (1999) *Correspondence*, ed. by Zweig, A. Cambridge University Press, UK.

Kant, I. (2000), *Critique of the Power of Judgment*, ed. by P. Guyer. Cambridge University Press, UK.

Kant, I. (2002), *Critique of Practical Reason*, ed. by W. S. Pluhar, intr. by S. Engstrom. Cambridge University Press, UK.

Kant, I. (2007), *Anthropology, history, and education*, ed. by Zöllner, G., & Louden, R. B., Cambridge University Press, UK.

Kant, I. (2011), *Die Religion innerhalb der Grenzen der blossen Vernunft*, ed. by Höffe, O. Walter de Gruyter.

Kant, I. (2012), *Lectures on Anthropology*. Cambridge University Press.

Kant, I. (2014), *Groundwork of the Metaphysics of Morals: A German–English Edition*, ed. by Gregor, M., & Timmermann, J. Cambridge University Press.

Kant, I. (2017), *The Metaphysics of Morals*, ed. by Denis, L., & Gregor, M. J. Cambridge, United Kingdom: Cambridge University Press.

Kant, I. (2019), *Religion within the boundaries of mere reason and other writings*, ed. by Wood, A. W., Giovanni, G. D., & Adams, R. M., Cambridge University Press, UK.

Korsgaard, C. M. (1996), *Creating the kingdom of Ends*, Cambridge University Press, UK.

Madore, J. (2011), *Difficult Freedom and Radical Evil in Kant: Deceiving Reason*. Bloomsbury Publishing, UK.

McMullin, I. (2013), Kant on Radical Evil and the Origin of Moral Responsibility, *Kantian Review*, no. 18(1), pp. 49 – 72.

- Mele, A. (1983), "Self-Deception" *Philosophical Quarterly*, no. 33, pp. 365 – 377.
- Mele, A., (2001), *Self-Deception Unmasked*, Princeton University Press, US.
- Mijović-Prelec, D., & Prelec, D. (2010), "Self-deception as self-signalling: a model and experimental evidence", *Philosophical Transactions of the Royal Society B: Biological Sciences*, no. 365(1538), pp. 227-240.
- Pasternack, L. (1999), "Can self-deception explain akrasia in Kant's theory of moral agency?" *Southwest Philosophy Review*, no. 15(1), pp. 87 – 97.
- Pasternack, L. (2013), *Routledge philosophy guidebook to Kant on Religion within the boundaries of mere reason*, Routledge, UK.
- Papish, L. (2018), *Kant on evil, self-deception, and moral reform*, Oxford University Press, UK.
- Paton, H. J. (1971), *The categorical imperative: A study in Kant's moral philosophy*, University of Pennsylvania Press, US.
- Pinheiro Walla, A. (2013), "Virtue and Prudence in a Footnote of the Metaphysics of Morals (MS VI: 433n)", *Jahrbuch für Recht Und Ethik/Annual Review of Law and Ethics*, no. 21, pp. 307 – 322.
- Sensen, O. (2011), "Kant's conception of human dignity", *Kant-Studien*, no. 100, pp. 174-212.
- Sticker, M. (2017), "When the Reflective Watch-Dog Barks: Conscience and Self-Deception in Kant", *The Journal of Value Inquiry*, no. 51(1), pp. 85 – 104.
- Sticker, M. (2021), *Rationalizing (Vernünfteln)*, Cambridge University Press, Cambridge.
- Szaif, J., & Thurnherr, U. (2004), "Wahrhaftigkeit", in J. Ritter, K. Gründer & G. Gabriel (Hg.), *Historisches Wörterbuch der Philosophie*, no. 12, 42. Schwabe Verlag, Basel.
- Timmermann, J., & Gregor, M. (2011), *Immanuel Kant groundwork of the metaphysics of morals*, Cambridge University Press, UK.
- Timmermann, J. (2000), Kant und die Lüge aus Pflicht. *Philosophisches Jahrbuch*, no. 107, pp. 267 - 283.
- Van Leeuwen, N. (2013), "Self-Deception", in H. LaFollette (ed.), *International Encyclopedia of Ethics*, Wiley-Blackwell, US.
- Wehofsits, A. (2020), "Passions: Kant's psychology of self-deception", *Inquiry*, pp. 1 – 25.
- Wood, A. W. (1970), *Kant's moral Religion*, Cornell University Press, Ithaca.

Wood, A. W. (1999), *Kant's ethical thought*, Cambridge University Press, UK.

