UNIVERSITÄT BAYREUTH

Revisiting Self-Deception: the Relevance of Kant's Moral Psychology for Contemporary Challenges in Environmental Ethics

Cumulative Dissertation submitted by

Maria Eugênia Zanchet Bordignon in 15th May 2023

to obtain the degree of doctor philosophiae

Department of Philosophy

Kulturwissenschaftliche Fakultät

Universität Bayreuth

Reviewers:

Dr. Alice Pinheiro Walla, McMaster University, Canada

Dr. Martin Sticker, University of Bristol, United Kingdom

Dr. Bryan Lueck, Southern Illinois University Edwardsville, United States

The inauthentic word (...) results from the dichotomy
that is established between its constituent elements.
(...) It is a hollow word, from which
no indictment of the world may be hoped for.

— Paulo Freire, Pedagogy of the Oppressed, p. 108

Revisiting Self-Deception:
The Relevance of Kant's Moral Psychology
for Contemporary Challenges in Environmental Ethics

Abstract: In recent years the concept of self-deception has been heavily discussed in Kant scholarship. At the same time, topics on environmental ethics, such as greenwashing and animal welfare are rapidly gathering momentum. However, despite the centrality of self-deception in Kant's ethical theory and its potential applicability to green issues, analytical studies on the intersection between the two subjects are virtually nonexistent. To challenge the underlying assumption that Kant's practical philosophy provides little room for addressing environmental issues, it is imperative that further explorations of self-deception be more closely applied to such contemporary challenges. The present dissertation aims to address this need by offering an exploration of the relevance of Kant's account of moral psychology to two topics in environmental ethics: the individual adherence of symbolic green narratives (SGNs) and animal welfare. I contend that, with no substantial revisionism, Kant's approach to self-deception remains relevant to these moral problems, as it is not only adequate but also a requisite for the efficacy of his practical philosophy in addressing the ethical issues under analysis. Intending to demonstrate the relevance and centrality of self-deception in Kant, this cumulative dissertation is composed of three articles: "Towards a Holistic View of Self-Deception in Kant's Moral Psychology" (A1) explores the concept of self-deception in Kant's practical theory by drawing on the parallel between this phenomenon and external lying; "The Moral Psychology of Individual-level Adherence to Symbolic Green Narratives" (A2) demonstrates the efficacy of a Kantian view of self-deception to bridge the explanatory gap allowed by contemporary theories of economics and social psychology; and "Regarding Animals: Kant's Account of Self-Deception and Its Relevance to Animal Welfare Advocacy" (A3) establishes the essential role of self-deception when it comes to making Kant's practical philosophy relevant to animal welfare. In den letzten Jahren ist das Konzept der Selbsttäuschung in der Kantforschung stark diskutiert worden. Gleichzeitig haben Themen der Umweltethik, wie Greenwashing und Tierschutz, an Bedeutung gewonnen. Trotz der zentralen Bedeutung der Selbsttäuschung in Kants ethischer Theorie und ihrer potenziellen Anwendbarkeit auf Umweltthemen gibt es jedoch so gut wie keine analytischen Studien zu den Überschneidungen zwischen diesen beiden Themen. Um die zugrundeliegende Annahme zu widerlegen, dass Kants praktische Philosophie wenig Raum für die Behandlung von Umweltthemen bietet, ist es unerlässlich, dass weitere Untersuchungen zur Selbsttäuschung stärker auf solche zeitgenössischen Herausforderungen

angewandt werden. Die vorliegende Dissertation zielt darauf ab, dieses Bedürfnis zu befriedigen, indem sie eine Untersuchung der Relevanz von Kants Darstellung der moralischen Psychologie für zwei Themen der Umweltethik anbietet: die individuelle Befolgung symbolischer grüner Narrative (SGNs) und den Tierschutz. Ich behaupte, dass Kants Ansatz zur Selbsttäuschung ohne wesentliche Revisionen für diese moralischen Probleme relevant bleibt, da er nicht nur angemessen, sondern auch eine Voraussetzung für die Wirksamkeit seiner praktischen Philosophie bei der Behandlung der untersuchten ethischen Fragen ist. Um die Relevanz und Zentralität der Selbsttäuschung bei Kant aufzuzeigen, besteht diese kumulative Dissertation aus drei Artikeln: "Towards a Holistic View of Self-Deception in Kant's Moral Psychology" (A1) untersucht das Konzept der Selbsttäuschung in Kants praktischer Theorie, indem es die Parallele zwischen diesem Phänomen und der externen Lüge heranzieht; "The Moral Psychology of Individual-level Adherence to Symbolic Green Narratives" (A2) zeigt, dass eine kantische Sichtweise der Selbsttäuschung ausreicht, um die Erklärungslücke zu schließen, die die zeitgenössischen Theorien der Wirtschafts- und Sozialpsychologie erlauben; und "Regarding Animals: Kant's Account of Self-Deception and Its Relevance to Animal Welfare Advocacy" (A3) legt die wesentliche Rolle der Selbsttäuschung dar, wenn es darum geht, Kants praktische Philosophie für den Tierschutz relevant zu machen.

Table of Contents

List of Abbreviations

Citations from Kant's works are provided alongside their corresponding volume in the *Gesammelte Schriften*, Ausgabe der Preußischen Akademie der Wissenschaften. The translations are sourced from *The Cambridge Edition of the Works of Immanuel Kant*. For individual works, the following abbreviations will be used:

RL	Über ein vermeintes Recht aus Menschenliebe zu lügen (1797), On a Supposed Right to Lie from Philanthropy	8:423-430
MS	Metaphysik der Sitten (1797-1798), Metaphysics of Morals	6:203-494
REL	Religion innerhalb der Grenzen der bloßen Vernunft (1793 - 1794), Religion within the Boundaries of mere Reason	6:1-202
PÄD	Pädagogik (1803), Lectures on Pedagogy	9:437-499
ANTH	Anthropologie in pragmatischer Hinsicht (1798), Anthropology from a Pragmatic Standpoint.	7:117-334
CORR	Kant’s Briefwechsel (1749 - 1800), Correspondence	10:07-12:268
KpV	Kritik der praktischen Vernunft (1788), Critique of Practical Reason	5:1-164
VE	Vorlesungen über Ethik (1760 - 1794), Lectures on Ethics	27:1-732
GMS	Grundlegung zur Metaphysik der Sitten (1785), Groundwork of the Metaphysics of Morals	4:385-464

1. Motivation, Hypothesis and Aims

1.1. Introduction, Research Question, and General Aims

It is not particularly surprising that, in contexts where lifestyles and decision-making typically revolve around social appearances, the ideas of deception and self-deception often arise[1]. In such settings, self-deception may be construed as a symptom of Western contemporary ways of life, in which it is common to ground decisions on the appearances of things around us, on the appearance of our (often merely performative) actions, as well as on the appearances we want to convey to our peers and to ourselves[2].

Yet there are plenty of compelling reasons to assume that self-deception is hardly unique to our times. Immanuel Kant had not only focused his attention on the phenomenon, but also viewed it as a trait of human nature. Self-deception significatively pervades his moral philosophy[3], playing a role in his notorious condemnation of lying[4,5] as well as in the well-known maxim "know

[1] Cf. Bloomfield, 2013; Froehlich, 2019; Enblad & Öhlander, 2019; Dong et al, 2019; Addawood et al., 2019; Dumas et al., 2020; Kolesnyk, et al., 2021. Generally, self-deception involves primarily social elements. On this matter, see also Dings, 2017.

[2] These may be read as hallmarks of a social phenomenon termed by Debord (2012) as *the spectacle*. It entails a monopolization of appearances of the states of affairs in the world, which are in turn ultimately grounded in mere abstractions. Debord puts forward the claim that once within an economic system where appearances are hogged, individuals become unable to act on the very essence of things. The topics discussed in this dissertation are consistent with Debord's view. The greenwashing performed by certain brands and institutions as well as the obfuscation of the eventual rationale behind individuals' adherence to such narratives (often an attempt to signify virtue) are examples of how economic or moral motivations may work to the advantage of mere appearances. Moreover, the strikingly successful practice of disassociating animals from the products made from their bodies and secretions - a phenomenon discussed by Adams (2015) as a discursive *absent referent* - similarly typifies the erasure described by Debord.

[3] The relevance of self-deception being such that it leads authors as Pasternack (1999, p. 92) or Reath (1989, 298–299) to regard self-deception as a condition for the possibility of immorality as such, as a condition of the highest degree of evil (Allison, 1996, p. 175; 1995, p. 148), as a condition of one's evil disposition (Papish, 2018, pp. *87–115*), as playing "an absolutely crucial role at *all levels* [*of evil*]" (Allison, p. 178, my emphasis), to mention some notable positions.

[4] MS 6:429 - 431. Overall, the account rendered in such passages provides the foundation for commentators' interpretation of self-deception as a form of internal lying. For an influential reading along these lines, see Wood (2009, p. 148).

[5] In addition to the MS, Kant's view of lying is also to be found in RL; ANTH 7:153; VE 7:605, 700 - 702. For a comprehensive debate concerning the three senses in which Kant uses the concept of lying, namely, juristic, ethical, and in the sense of right, see Mahon (2009).

thyself"[6], a hallmark of the efforts of modern philosophy[7]. For Kant, the duty against lying is also to be regarded as the duty against self-deception, as the latter is itself a form of lying (MS 6:429). As for the maxim "know thyself", it not only emphasizes one's capacities and limitations towards pure practical reason, but it crucially refers to the effects of a lack of truthfulness with regard to the agent's assessment of their own maxims and moral character (MS 6:441).

However, in contrast to Kant's systematic treatment of lying and his insistence on the importance of moral self-knowledge, self-deception does not seem to receive the same attention[8], as Kant's use of that concept is both widely dispersed[9] and applied to varied contexts (Papish, 2018, p. 69)[10]. One outcome of this setting is that self-deception still amounts to an overlooked component within Kant's canon (Sticker, 2021, p. 10), and only in the last few decades have philosophical analyses on self-deception emerged in Kant scholarship[11]. Even if Kant himself

---

[6] Cf. MS 6:441. This maxim stems from the philosophical writings of thinkers such as Plato, Socrates and Sophocles, later emerging in modern investigations about the nature of the self, led by philosophers such as Hume, Descartes, Locke and Kant.

[7] Claims revolving around self-deception are also to be found in Kant's account of the evil within his Religion (REL 6:38; 6:42f; 6:173-4; 6:200), arguably in his views on *dissimulation,* elaborated in his Pedagogy (PÄD 9:486) as well as in the exploration of the deceptive character of the inclinations in the Anthropology (ANTH 7:152).

[8] In addition, Papish (2018, pp. 4; 68) notes that in the literature, the concept of self-deception has long been assumed to be "self-explanatory".

[9] For example, in *The Metaphysics of Morals* (MS 6:429) self-deception is typified as a subcategory of deception in general. In *Anthropology from a Pragmatic Point of View* (ANTH 7:219) Kant makes a characterization of self-deception as entailing one's misleading attitude towards oneself when construing habits and inclinations as objective commands; in *Religion Within the Boundaries of Mere Reason*, self-deception is typically associated with Kant's theory of evil (REL 6:38, 642f), while simultaneously standing in his account of the moral principle of religious practice (REL 6:173 - 174) or his discussion of idolatry (REL 6:198 - 199).

[10] It is also worth noting that, as it amounts to an activity involving one's misuse of practical rationality as well as one's interest in providing moral justifications for one's own actions, self-deception is often treated as closely bound with rationalization [*Vernünfteln*], which most holistic and comprehensive treatment is given by Sticker (2021, pp. 08 -10; 2022, pp. 09 -13). This represents a second difficulty that comes along with the concept of self-deception, for it is not clear what the relationship between the two is - whether self-deception would be an aspect of rationalization or vice versa, or whether they would be interchangeable terms.

[11] For example, Baron (1988, p. 443 - 444) adopts the approach of examining self-deception as a process, and contends that the problem with self-deception lies in the agent's refusal to self-examination. Wood (2009) places self-deception alongside *akrasia* as a central component for the intelligibility of evil. Formosa (2009) emphasizes the importance of self-deception in making justificatory narratives plausible. In conjunction with passions, self-deception serves a major part in the dynamics of moral transgression; this is argued by Wehofsits (2020), who describes the workings of impassioned self-deception in the practice of immoral actions and Kant's motivations for the sharp rejection of passions. Potter (2002, pp. 384 - 386) characterizes the different forms of self-deception in Kant in terms of their objects and the possibility of reversibility. According to him, self-deception is the chief enemy of virtue and the main form of moral evil (from which one is never able to completely recover). In the same spirit, Moran (2014, p. 439) describes self-deception as a mechanism by which the agent falsifies one's own autonomy and "disregards the real unconditional practical principle because she has constructed a comfortable and convincing substitute based on her own inclination" to ultimately misrepresent one's transgressions in a misguided way. The most focused and influential treatments of self-deception have recently emerged. They have been undertaken by Papish (2018), who presents a comprehensive account of how we should conceive of self-deception and its relation with radical evil; and Sticker (2015; 2017; but especially 2021 and 2022) who in addition to providing an exegetical

fails to devise a systematic account of self-deception, the prominence he assigns to the phenomenon throughout his work[12] remains indisputable.

Meanwhile as the literature on self-deception is gathering momentum, topics that in Kant's time were considered philosophically fruitless or lacking in moral import[13] are now becoming increasingly relevant. In particular, it has only been in the last decade that topics of environmental ethics and animal matters have come into wide recognition as subjects deserving concentrated[14] studies.

However, in light of the climate crisis and the special attention being devoted to issues of environmental ethics[15], many assume that Kant's ethical theory would have little to offer on such topics. This contention is premised on three assumptions: (1) an overemphasis on the *anthropocentrism* of Kant's ethics, which in turn, (2) would hardly extend beyond the *indirect* duties that (3) *individual* moral agents have toward non-rational beings, such as animals or the environment. For instance, Gudorf & Huchingson's (2010, pp. 06 - 11) description of Kant's deontology exemplifies a characterization of Kant's ethical thought as overly anthropocentric[16].

---

examination of the concept of rationalization, also explores its workings and the role of the agent's concern for morality when it comes to moral transgressions and forms of pseudo-rationality.

[12] See for instance the passage where Kant describes self-deception as "blue dust" one throws in his own eyes (REL 6:38), as a phenomenon whereby one "*lie[s] to oneself* in the interpretation of the moral law to its prejudice" (REL 6:42f, *my emphasis*), or even as associated with his account of evil, as previously observed (*cf. note 01*). Along similar lines, Kant's economy of duties also typifies the prevalence of self-deception within his ethics, as explored in this dissertation.

[13] As with his contemporaries, Kant himself saw no moral problem in eating meat; this becomes evident in his discussion of fantastic virtue (MS 6:409), in which he suggests that turning "petty details" into moral decisions such as "whether I eat meat or fish, drink beer or wine" is a trait of a defective moral life.

[14] It is worth noting that the moral standing of animals was subject of discussions in ethics within certain ancient traditions, thus informing thoughts that came to ground contemporary legal systems (Rollin, 2017, p. 09). However, such reflections on animal matters were generally parasitic on other topics, or often superficial, as exemplified by animal rights discussions of the early modern period (Garret, 2011). Kant himself has written only sparsely on animal matters, despite the salience of the idea of animality within his ethics (Baumeister, 2022). Conversely, environmental ethics has only grown into a major issue in the past few decades, driven by practical matters of climate emergency and by what is currently discussed under the term *climate crisis*.

[15] Not until the early 1980s, with the institutionalization of environmental education organizations and other non-profit organizations, did topics related to environmental education become progressively mainstream. The 1987 UN report *Our Common Future* illustrates this tendency. This report not only moves towards making public the evidence concerning the effects of human action in its handling of finite natural resources, but also towards a pedagogical approach, providing insights that can be seen nowadays as a means to address the issue on how we should, as humanity, behave in the face of environmental degradation.

[16] Although it is important to note that they later make the case for thinkers like John Passmore, who argue for the philosophical competence of the anthropocentric perspective in addressing environmental issues. Critiques of Kant's anthropocentrism can be found in Singer (2002, especially p. 203), Chakrabarty (2016) and Reagan (2004).

As for (2), to depict the diversity of obligations allowed by indirect duties, Hooley & Nobis (p. 23) invite the reader to picture a "meat-loving environmentalist" who, instead of reducing meat consumption, takes alternative actions that minimize his environmental damage, a maneuver that is only possible because in Kantian ethics the welfare of non-rational beings is “merely” an indirect duty. Whereas concerning (3), characterizations of Kant's ethics as individualistic have prompted authors such as Jamieson (2014) to denounce it as unfit to deal with environmental challenges that require collective and inter-generational solutions.

The alleged incompatibility between Kant and environmental ethics has been recently challenged, whether through reconciling Kant's practical philosophy with issues in environmental ethics[17] or by revisions of Kant's moral thought brought about by authors such as Korsgaard (2012; 2013; 2018), Vereb (2019) and Möller (2022). However, the field where environmental ethics and Kant's views about self-deception meet remains largely unexplored. Thus, addressing the question as to *whether self-deception may be relevant to the theoretical developments on green topics* becomes imperative.

The overall aim of this dissertation is to advance the debate about self-deception not solely within Kant scholarship, but also to broaden the application of Kant's ethics to ethical challenges of contemporary Western societies. This dissertation's desideratum is to explore the concept of self-deception in Kant's ethical system, in order to outline the relevance of his perspective for dealing with certain contemporary ethical phenomena within the field of environmental ethics, such as the adherence to deceptive green narratives or the rejection of animal cruelty.

The thesis I put forward in this dissertation concerns the ascription of two features to Kant's account of self-deception and his ethics when it comes to making sense of the following aspects: adequacy and efficacy. My hypothesis is that without substantial revisionism, Kant's approach to self-deception remains relevant for tackling certain moral problems, as it is not only *adequate* in

[17] Cf. Altman (2014); Wilson (2017); Wyrębska-Đermanović, (2021).

complementing further theoretical developments (A2), but is also a requisite for the *efficacy* of his practical philosophy (A3) in addressing particular ethical issues.

This thesis will therefore demonstrate (A) the adequacy, i.e., the sufficiency of a Kantian view of self-deception in bridging the explanatory hiatus between contemporary theories of economics and social psychology; and (B) the essential role of self-deception for rendering Kant's practical philosophy relevant to animal welfare. In turn, these outcomes will contribute to Kant scholarship, by establishing the centrality of self-deception in Kant's ethics.

By using Kant's account of self-deception to explicate individual adherence to green narratives and its interplay with the practical identities of certain agents, article A2 demonstrates the *adequacy* of Kant's ethics when attempting to make sense of a phenomenon that has been only incompletely described by mainstream theories in economics and social psychology. This discussion simultaneously adds to the debate held in article A3 on the centrality of the duty against self-deception for rendering Kant's ethics philosophically *efficacious* for cases of applied ethics, such as the human treatment towards non-human animals (hereafter, animals). Both critical analyses bring forth insights into the importance of self-deception not only for Kant's ethical system but for the relevance of Kant's ethics in the face of contemporary moral problems and benefit from the prior investigation accomplished in article A1.

1.2. Specific Aims

The specific aim of A1 is to explore the concept of self-deception in Kant from the parallel he draws in §9 of the *Doctrine of Virtue* between self-deception and external lying and, building on this analysis, to explain why Kant objected to self-deception so sharply.

Kant's account of self-deception as a form of (external) lying has received a fair amount of criticism in the literature. For example, Papish (2018, p. 70) claims that mirroring self-deception on external lying amounts to a "ill- conceived framework for understanding self-deception" and

that, although Kant seems to support this framework, "he is aware of the ways that self-deception cannot be thought of as a form of deception more generally". Analogously, Sticker suggests that mirroring self-deception on external lying (that results from Kant's attempt to make sense of an apparently paradoxical phenomenon) may be equally problematic. After all, Sticker explains (2021, p. 35), self-deception requires that "one cognitive system of an agent employs a deceptive strategy against another system of the same agent and unbeknownst to that latter system, much like when one agent deceives another". The way Kant accounts for this bifurcation is, according to Sticker, troublesome as "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" (Zanchet, 2022a, p. 196).

The discussion held in A1 concentrates on complementing Papish and Sticker's analyses by showing that a further investigation of the parallel between self-deception and external lying may provide a more holistic view of the phenomenon. Through examining concepts such as personality, dignity, and humanity, my approach addresses Kant's potential motivations for framing this parallel in the first place. Notwithstanding Kant's myriad mentions of these concepts in the context of deception[18], the current literature on the phenomenon seems to overlook their role in the framework of self-deception. For example, Papish (2018, p. 59) mentions the concepts of humanity and personality in the context of the predispositions discussed by Kant in Religion, in a very brief reconstruction of Muchnik's claims. Likewise, Sticker (2021, p. 19) succinctly alludes to dignity as part of his argument that self-deception transforms and inevitably corrupts the moral law, thus depriving the agent of one's dignity. Overall, the literature considers the meaning of these terms as self-evident, thus failing to address the counterintuitive claim that it is possible for a human being to lose one's dignity or personality. In my approach these concepts are considered as a primary aspect of Kant's framework of self-deception, as they are closely connected to the moral implications of the wrongdoing entailed by both forms of lying. My analysis fulfills a theoretical lacuna towards a more holistic understanding of the phenomenon of self-deception in Kant.

[18] Cf. MS 6:420, 423, 425, 429, 436, 446, 462; CORR 11:332; PÄD 9:488-9; GMS 4: 405. KpV 5:71. See also ANTH 7:127.

As previously mentioned, the aim of A1 is to examine the rationale underlying Kant's stringent objections against external lying and self-deception. This will entail overcoming a theoretical tension that might emerge as a result of over-stressing the parallel between the two forms of lying. I argue that "modeling self-deception on external lying adds up to the premises of what I (...) describe as *the teleological claim about self-deception*" (Zanchet, 2022a, p. 197) and that one must be cautious not to jump to the view that Kant buys into teleological claims to account for the transgression implied by self-deception.

The assessment of self-deception carried out in A1 establishes that the key point of self-deception does not merely consist in the misuse of rational faculties, but rather in a threat to the very core of these capacities, as both forms of deception impinge on the way *one regards oneself* while simultaneously twisting the agent's view of one's relation to pure practical reason.

Having that analysis in hand, further applications of Kant's framework of self-deception to two contemporary problems follow, i.e., adherence to deceptive green narratives (A2) and the issue of animal welfare (A3).

A2 offers a Kantian-inspired analysis of the phenomenon of individual adherence to misleading green narratives, among which the contemporary problem encapsulated within the idea of *greenwashing* lies.

In A2, I provide a framework that accounts for the success of symbolic green narratives (SGN) in misleading their audience - notwithstanding the fact that agents acknowledge them as deceptive. For this purpose, I argue that contemporary theories of economics and social psychology often fail to adequately address the complexity of the phenomenon, which involves cognitive, economic, but most importantly, *moral self-regarding* elements.

In the discussion held in A2, I reference the theories of Nyilasi et al. (2012) and Mitchell & Ramey (2011) as representing approaches concerning the effectiveness of SGNs in deceiving

their agents. While Nyilasi et al. explain that perceived falsity of these narratives gives rise to negative attitudes towards green deceivers, Mitchell & Ramey (2011) turn to social and anthropological aspects to account for the success of such narratives in misleading their audience. I advance these theories by arguing that while agents are often driven to endorse SGNs for social reasons, they are also likely to endorse them as a form of self-affirmation. I contend that behind the endorsement of such narratives are underlying psychological strategies such as self-deception, used by agents to bolster their practical identities.

Thus, the specific aim in A2 is to show that "in light of the perceived symbolic or deceptive aspect of [green] narratives, individuals' adherence is significantly influenced by first-person psychological, in addition to (third-person) anthropological and social aspects" (Zanchet, 2022 forthcoming) and, most importantly, to establish the centrality of the moral aspects in these agents' decision making when it comes to green issues.

Drawing on the Kantian philosopher Christine Korsgaard's first-person theory of morality and elements that underlie Kant's view of human moral psychology, I contend that agents' practical identities shape their commitments to green narratives in general. Furthermore, I show that self-deception plays a key role in ensuring that agents can hold certain practical identities while at the same time performing actions that violate the self-attribution of those same identities. Such a conclusion is only possible given the argumentative effort undertaken in the third section of A2, where I lay out the ethical component of such narratives. It is precisely by bringing such an ethical feature into focus that it is possible to draw a complementary framework to the aforementioned theories, as they fail to account for the self-referential attitudes implied in adherence to SGNs. This ethical component suggests that agents engage in self-deception as an attempt to justify the moral transgressions that threaten their self ascription of certain practical identities. Insofar as one embraces certain conceptions of practical identity without at the same time being committed to actions consistent with that identity, a conflict results. Self-deception arises as an outcome of such a conflict, i.e., as a way to preserve one's moral integrity.

My analysis therefore shows that when construed upon self-deception, practical identities may ultimately play a detrimental part on the agent's practical reasoning. In the absence of the agent's concern when forming their judgments and accessing their maxims and actions, self-deception may operate side by side with one's practical identity so as to hinder the agent from being truthful about one's will, maxims, and character.

Similarly to A2, A3 also extends Kant's account of self-deception to a contemporary problem, namely, violations of animal welfare. Yet, in contrast to A2, A3 focuses directly on Kant's theoretical framework, addressing the epistemic complexities of the phenomenon of self-deception in order to examine its relevance for constructing a moral account that considers animal welfare as central.

It is within the argument pursued in A3 that the intent of this dissertation is made most clear, for the analysis undertaken in that article provides a response to the aforementioned criticisms. While within animal welfare movements prevails a view that Kant's ethics is overly anthropocentric, progressive social movements often charge theories such as Kant's of flattening "cultural nuances, [as] they leave little or no room for social and situational particularities that mark cultural diversity" (Zanchet, 2022b, p. 12). Both criticisms partake in the premise that Kant's theory, supposedly overly concerned with rational agents and their moral duties, and furthermore rooted in the idea of universal moral principles, would fail to accommodate non-rational beings such as animals or the environment as well as the social aspects that shape cultural diversity.

Accordingly, A3 pursues two specific aims. Primarily, A3 shows that adopting Kant's framework of self-deception is adequate for the purposes of animal advocacy. To this end, I explore the limitations of Kant's ethical theory, the main one stemming from the epistemic factors entailed by the concept of self-deception, which in turn pose a challenge to the relevance of Kant's approach in this matter. I then contend that the duty against self-deception constitutes a "back-up

duty" so that other duties in Kant's ethical system can be accommodated in a theory suitable to protect animals from human harm.

I show in A3 that even without going through revisions challenging the well-established view that Kant does not grant animals direct duties, further elements may count as insightful resources in the advocacy for animal welfare. My reconstruction of the self-regarding duties to strive for moral perfection and self-knowledge reinforces this point. Within this framework, I argue that the duty against self-deception complements these aforementioned duties, for it ensures that the agent's assessment of one's own compliance with them depends on an attitude of truthfulness.

The analysis of the merits of Kant's framework of self-deception will support my argument that the duty against self-deception amounts to

> a fundamental step one should take in order to avoid performing actions of animal exploitation especially when it comes to the main forms of animal abuse discussed throughout the empirical literature (Zanchet, 2022b, p. 13)

The second specific aim of A3 is to debunk the criticisms against Kant's theory concerning its relevance for making the case for animal welfare. For this, I show that Kant's account of moral progress presupposes a connection between individual and collective progress and that, unlike the picture these criticisms paint, Kant's theoretical framework preserves the place for moral progress as a collectively sought end.

2. Conclusion

By exploring the concept of self-deception, this dissertation broadens the understanding of the phenomenon in Kant scholarship while at the same time demonstrating the concept's applicability as a means of making sense of contemporary moral problems within the field of environmental ethics.

As a result of the conducted investigations, this dissertation's research question, namely, *whether self-deception may be relevant to the theoretical developments on green topics*, is positively answered.

While lending salience to Kant's ethics vis-à-vis theories in economics and moral psychology, A2 meets its desideratum by filling the gap left by descriptive accounts concerning the mechanisms involved in an audience's adherence to narratives perceived as deceptive. My analysis underscores the moral and normative elements that explicate such adherence when it comes to deceptive self-regarding attitudes of moral self-affirmation.

Through adding a first-person normative layer to the debate and exploring the moral psychology behind mechanisms of pseudo-rational adherence to deceptive narratives, A2 shows that Kant's account of self-deception along with Korsgaard's theoretical framework on formation and preservation of practical identities provide effective explicatory pathways to address the examined phenomenon.

A3 establishes the *efficacy* of Kant's theoretical approach to the issue of animal welfare. This is achieved after an analysis that weighs the defensible aspects as well as the potential shortcomings of Kant's theory in accommodating animal welfare, and concludes that the key for rendering Kant's economy of duties effective is precisely the duty against self-deception, which functions as a form of back-up to the duties towards self-knowledge and moral self-perfection.

The successful applicability of Kant's ethics to the animal welfare problem is contingent upon his account of self-deception in a manner that also leads one to acknowledge the phenomenon's inherent complexity.

The analytical endeavors concerning the aforementioned ethical problems undertaken in this dissertation lead to the conclusion that even without incurring significant revisions, Kant's account of self-deception is not only adequate but is also a *conditio sine qua non* efficiently dealing with certain contemporary moral problems from a Kantian perspective.

It is worth mentioning that the overall goal of this dissertation, namely to expand the concept of self-deception within and beyond Kant scholarship, concurrently succeeds in helping to draw attention to the intersection between environmental issues and Kant's views on self-deception.

Examining green topics interdisciplinarily is important because it often provides novel forms of framing a problem, thus making possible solutions that were otherwise difficult to reach. For example, using Kant as an interpretive linchpin for addressing animal welfare may foster constructive and progressive shifts in the theoretical foundations of animal ethics, which are currently predominantly utilitarianist and face a variety of practical hindrances.

Moreover, readings that take into account first-person perspectives are similarly relevant, in a Kantian sense. This is because self-scrutiny and the individual pursuit of virtue are the shaping blocks of Kant's envisioned society, the Kingdom of Ends, where social and individual elements are enacted towards a common end (Herman, 1997).

The examinations undertaken in A2 and A3 in turn make up the yet limited theoretical framework of application of Kant's ethics to environmental issues, and may not only be of benefit for providing input for further developments, but also as a framework of analysis to be scaled up and expanded to other green topics. Overall, the matters of the present dissertation

further contribute to building momentum so that green topics can be placed more prominently in the agenda of Kant scholarship.

Finally, this dissertation demonstrates the value of delving into Kant's moral philosophy nowadays. In contrast to what is commonly assumed, Kant's theory has not been surpassed, and his remarks on moral psychology as well as his views on self-deception have proven to be not only in line with major tenets established by contemporary moral psychology, but also insightful contributions to understanding a number of contemporary problems. Rather than attempts to merely describe human behavior, Kant's framework lays out normative guidelines as to how to navigate the moral terrain. When dealing with questions such as whether we should endorse greenwashing, or whether we should participate in promoting violence towards animals, Kant's ethics and his account of self-deception are invaluable resources of both answers and challenges.

3. Bibliography

Addawood, A., Badawy, A., Lerman, K., & Ferrara, E. (2019). Linguistic cues to deception: Identifying political trolls on social media. *Proceedings of the international AAAI conference on web and social media* (Vol. 13, pp. 15-25).

Altman, M. C. (2014). Animal Suffering and Moral Character. Kant and Applied Ethics.*The Uses and Limits of Kant's Practical Philosophy*, 13-44.

Allison, H. E. (1995). Reflections on the banality of (radical) evil: A Kantian analysis. *Graduate Faculty philosophy journal*, *18*(2), 141-158.

Allison, H. E. (1996). *Idealism and freedom: essays on Kant's theoretical and practical philosophy*. Cambridge University Press.

Baumeister, D. (2022). *Kant on the Human Animal: Anthropology, Ethics, Race*. Northwestern University Press.

Baron, Marcia (1988). What is wrong with self-deception? In Brian P. McLaughlin & Amelie O. Rorty (eds.), *Perspectives on Self-Deception*. University of California Press. pp. 431- 449.

Bloomfield, P. (2013). Social Media, Self-Deception, and Self-Respect. *Social Media and the Value of Truth*, 28-39.

Chakrabarty, D. (2016). Humanities in the Anthropocene: The crisis of an enduring Kantian fable. *New Literary History*, *47*(2), 377-397.

Debord, G. (2012). *Society of the Spectacle.* Bread and Circuses Publishing.

Dings, R. (2017). Social strategies in self-deception. *New Ideas in Psychology*, *47*, 16-23.

Dong, M., van Prooijen, J. W., & van Lange, P. A. (2019). Self-enhancement in moral hypocrisy: Moral superiority and moral identity are about better appearances. *PloS one*, *14*(7), e0219382.

Dumas, T. M., Maxwell-Smith, M. A., Tremblay, P. F., Litt, D. M., & Ellis, W. (2020). Gaining likes, but at what cost? Longitudinal relations between young adults' deceptive like-seeking on instagram, peer belonging and self-esteem. *Computers in Human Behavior*, *112*, 106467.

Enblad, L., & Öhlander, E. (2019). Deception and Self Deception : An investigation of Multi-level marketing distributors and their deceptive practices on social media (*Dissertation*).

Formosa, P. (2009). Kant on the Limits of Human Evil. *Journal of Philosophical Research*, 34: 189–214.

Froehlich, T. J. (2019). The role of pseudo-cognitive authorities and self-deception in the dissemination of fake news. *Open Information Science*, *3*(1), 115-136.

Garret, A. (2011). Animals and Ethics in the History of Modern Philosophy. In: Tom L. Beauchamp and R. G. Frey (Eds.) *The Oxford Handbook of Animal Ethics*. Oxford University Press.Adams, C. J. (2015). *The sexual politics of meat: A feminist-vegetarian critical theory.* Bloomsbury Publishing USA.

Gudorf, C. E., & Huchingson, J. E. (2010). *Boundaries: A casebook in environmental ethics*. Georgetown University Press.

Herman, B. (1997). A Cosmopolitan Kingdom of Ends. In Andrew Reath, Barbara Herman, Christine Korsgaard, (eds.) *Reclaiming the History of Ethics. Essays for John Rawls*. Cambridge: Cambridge University Press, pp. 187 - 213.

Hooley, D., & Nobis, N. (2022). Eating. *The Routledge Companion to Environmental Ethics* (pp. 20-30). Routledge.

Jamieson, D. (2014). *Reason in a Dark Time: Why the Struggle Against Climate Change Failed - and What it Means for our Future*. Oxford: Oxford University Press.

Kant, I. (1798). Anthropology from a Pragmatic Point of View, Zöller, G., & Louden, R. B. *Anthropology, History, and Education* (2007). Cambridge, United Kingdom: Cambridge University Press.

_____. (1803). Lectures on Pedagogy, In Zöller, G., & Louden, R. B., *Anthropology, History, and Education* (2007). Cambridge, United Kingdom: Cambridge University Press.

_____. (1996). On a Supposed Right to Lie from Philanthropy, in Gregor, M. J. (ed.) *Practical Philosophy*. Cambridge, United Kingdom: Cambridge University Press.

_____. (1997). *Lectures on Ethics*, Heath, P. & Scheenwind, J.B (Eds). Cambridge, United Kingdom: Cambridge University Press.

_____.. (1999). *Correspondence* (Vol. 11). Cambridge, United Kingdom: Cambridge University Press.

_____. (2011). *Immanuel Kant: Groundwork of the Metaphysics of Morals: a German–English Edition*. Cambridge, United Kingdom: Cambridge University Press.

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

_____. (2019). Religion Within the Boundaries of Mere Reason and other writings, ed. by Wood, A. W., Giovanni, G. D., & Adams, R. M. Cambridge, United Kingdom: Cambridge University Press.

Kolesnyk, D., de Jong, M. G., & Pieters, R. (2021). Gender gaps in deceptive self-presentation on social-media platforms vary with gender equality: A multinational investigation. *Psychological science*, *32*(12), 1952-1964.

Korsgaard, C. M. (2012). *A Kantian case for animal rights.*

_____. (2013). Personhood, animals, and the law. *Think*, 12(34), 25-32.

_____. (2018). *Fellow creatures: Our obligations to the other animals*. Oxford University Press.

Mahon, J. E. (2009). The Truth about Kant on Lies. *The Philosophy of Deception*, 201-224.

Mitchell, L., & Ramey, W. (2011). Look how green I am! An individual-level explanation for greenwashing. *Journal of Applied Business and Economics*, *12*(6), pp. 40 - 45.

Moran, K. (2014). Delusions of virtue: Kant on self-conceit. *Kantian Review*, *19*(3), 419-447.

Müller, N. D. (2022). *Kantianism for Animals: A Radical Kantian Animal Ethic*. Springer Nature.

Nyilasy, G., Gangadharbatla, H., & Paladino, A. (2012). Greenwashing: A consumer perspective. *Economics & Sociology*, *5*(2), 116.

Papish, L. (2018). *Kant on evil, self-deception, and moral reform*. Oxford University Press.

Pasternack, L. (1999). Can self-deception explain akrasia in Kant's theory of moral agency? *Southwest Philosophy Review*, *15*(1), 87-97.

Potter, N. (2002) Duties to Oneself, Motivational Internalism, and Self-Deception in Kant's Ethics. Mark Timmons (ed.), *Kant's Metaphysics of Morals: Interpretative Essays*. Clarendon Press.

Regan, Tom (2004). *The Case for Animal Rights. Berkeley*, CA: University of California Press.

Rollin, B. E. (2017). *A new basis for animal ethics: Telos and common sense*. University of Missouri Press.

Singer, P. (2002). *Animal Liberation*. New York: Ecco: HarperCollins Publishers.

Sticker, M. (2015). Educating the Common Agent. Kant on the Varieties of Moral Education. *Archiv für Geschichte der Philosophie* 97: 358–87

_____. (2017). When the reflective watch-dog barks: conscience and self-deception in Kant. *The Journal of Value Inquiry*, *51*, 85-104.

_____. (2021). *Rationalizing (Vernünfteln)*. Cambridge University Press.

_____. (2022). Who is Rationalising? On an Overlooked Problem for Kant's Moral *Psychology and Method of Ethics. Kantian Journal*, 41(1).

Vereb, Z. T. (2019). *The Case for the Green Kant: A Defense and Application of a Kantian Approach to Environmental Ethics*. University of South Florida.

Wehofsits, A. (2020). Passions: Kant's psychology of self-deception. *Inquiry*, 1-25.

Wilson, H. L. (2017). The green Kant: Kant's treatment of animals. *Food Ethics*, 5-13.

Wood, A. (2009). Kant and the Intelligibility of Evil. S. Anderson-Gold & P. Muchnik (Eds.), *Kant's Anatomy of Evil (pp. 144-172)*. Cambridge: Cambridge University Press.

World Commission on Environment and Development (1987). *Our Common Future*. Oxford: Oxford University Press.

Wyrębska-Đermanović, E. (2021). Kantian moral education for the future of humanity: The climate change challenge. *Journal of Philosophy of Education*, 55(6), 1045-1056.

Zanchet, M. E. (2022a). Towards a Holistic View of Self-Deception in Kant's Moral Psychology. *Con-Textos Kantianos. International Journal of Philosophy*, (16), 194-219.

_____. (2022b). Regarding Animals: Kant's Account of Self-Deception and Its Relevance to Animal Welfare Advocacy. *Aufklärung,* v.9, n.3, pp. 11 - 30

_____. (2023, forthcoming) The Moral Psychology of Individual-level Adherence to Symbolic Green Narratives. *Estudos Kantianos.*

4. Appendix

4.1. A1 - Towards a Holistic View of Self-Deception in Kant's Moral Psychology

Zanchet, M. E. (2022). Towards a Holistic View of Self-Deception in Kant's Moral Psychology. *Con-Textos Kantianos. International Journal of Philosophy*, (16), 194-219

CON-TEXTOS KANTIANOS.
International Journal of Philosophy
N.º 16, December 2022, pp. 194-219
ISSN: 2386-7655
Doi: 10.5281/zenodo.7404175

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

[Received: June 9th 2022
Accepted: November 19th 2022]

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-deception[4], 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.

CON-TEXTOS KANTIANOS
International Journal of Philosophy
N.º 16, December 2022, pp. 194-219
027
ISSN: 2386-7655
Doi: 10.5281/zenodo.7404175

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 sort[6].

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".

CON-TEXTOS KANTIANOS
International Journal of Philosophy
N.º 16, December 2022, pp. 194-219
029
ISSN: 2386-7655
Doi: 10.5281/zenodo.7404175

the two "persons" Kant requires to make sense of self-deception[9] 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 means[10]. 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 beings[11], 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).

CON-TEXTOS KANTIANOS
International Journal of Philosophy
N.º 16, December 2022, pp. 194-219
031
ISSN: 2386-7655
Doi: 10.5281/zenodo.7404175

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 person[15]. 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 professes[16]. 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*.

CON-TEXTOS KANTIANOS
International Journal of Philosophy
N.º 16, December 2022, pp. 194-219
033
ISSN: 2386-7655
Doi: 10.5281/zenodo.7404175

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)

CON-TEXTOS KANTIANOS
International Journal of Philosophy
N.º 16, December 2022, pp. 194-219
035
ISSN: 2386-7655
Doi: 10.5281/zenodo.7404175

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 raised[19] 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 [*abwürdigen*] 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 nature[20] 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 choice[21]. 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.

CON-TEXTOS KANTIANOS
International Journal of Philosophy
N.º 16, December 2022, pp. 194-219
037
ISSN: 2386-7655
Doi: 10.5281/zenodo.7404175

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 dignity[26].

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".

CON-TEXTOS KANTIANOS
International Journal of Philosophy
N.º 16, December 2022, pp. 194-219
039
ISSN: 2386-7655
Doi: 10.5281/zenodo.7404175

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 true[29]. 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 lying[30] 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 truthfulness[31]. 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.

CON-TEXTOS KANTIANOS
International Journal of Philosophy
N.º 16, December 2022, pp. 194-219
041
ISSN: 2386-7655
Doi: 10.5281/zenodo.7404175

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 truthfulness[33]. 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 insomuch 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 imperative[34] 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).

CON-TEXTOS KANTIANOS
International Journal of Philosophy
N.º 16, December 2022, pp. 194-219
043
ISSN: 2386-7655
Doi: 10.5281/zenodo.7404175

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 themselves[36]. 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 interpreters[37] 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.

CON-TEXTOS KANTIANOS
International Journal of Philosophy
N.º 16, December 2022, pp. 194-219
045
ISSN: 2386-7655
Doi: 10.5281/zenodo.7404175

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), Zöller, 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 Zöller, 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.

CON-TEXTOS KANTIANOS
International Journal of Philosophy
N.º 16, December 2022, pp. 194-219
ISSN: 2386-7655
Doi: 10.5281/zenodo.7404175

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

4.2. A2 - The Moral Psychology of Individual-level Adherence to Symbolic Green Narratives

Zanchet, M. E. (2024) The Moral Psychology of Individual-level Adherence to Symbolic Green Narratives. *Estudos Kantianos, v. 12 n. 1*

THE MORAL PSYCHOLOGY OF INDIVIDUAL-LEVEL ADHERENCE TO SYMBOLIC GREEN NARRATIVES: A KANTIAN THEORETICAL APPROACH

Maria Eugênia Zanchet, Universität Bayreuth

INTRODUCTION

In recent years, theories have tried to account for public adherence to deceptive phenomena such as greenwashing. A number of these theories have sought to answer why people choose to take actions that endorse narratives they acknowledge to be merely symbolic and, to a large extent, intentionally manufactured to deceive them.

In the context of narratives targeting green or environmental issues, this phenomenon has been addressed by theories of economics and social psychology. By offering descriptions in terms of the motivational aspects that govern consumers' choices, such theories tend to overfocus on the individual's perception concerning these narratives' falsity, consequently neglecting the *moral* aspect of people's decisions. In addition, even theories engaged in addressing moral aspects tend to overlook further components that are key to a deeper understanding about why agents endorse green narratives despite recognizing that they are deceptive. As a result, the mechanisms at work in cases where green narratives succeed in misleading their audience remain opaque, leading to a partially obscured picture of the phenomenon.

My main goal in this paper is to explore the adherence to green narratives which the agent recognizes as deceptive or merely symbolic from a moral standpoint. I will show that once under a moral framework, we allow for an approach whereby the overlooked components become salient.

Two of these components will be explored. The first refers to the idea that human moral experience is essentially *normative*, which is particularly evident in situations that place the agent in a position to make decisions in light of certain principles. The second, in turn, concerns a feature that is fundamentally at play when such decisions are made, namely the *first-person perspective*. As a result, I will make the case for a way of approaching the phenomenon of green narratives from an ethical and normative standpoint which additionally focuses on self-regarding considerations.

In Section One, I elucidate what I mean by symbolic green narratives (SGN) and contextualize the problem from two theories that I believe represent the context I have just laid out. The first theory descriptively delineates the phenomenon, while the second one presents an attempt (albeit limited) to approach it from an ethical standpoint. I then discuss the limitations of these theories and argue that they should be complemented by a moral perspective approach.

Presenting this complementary theory will be my goal in the next two sections. To do so, I will draw from the theory of a Kantian philosopher credited with elaborating on the first-person perspective, Christine Korsgaard. A number of elements from Kant's moral psychology will be underlined to elucidate Korsgaard's theory, but also to provide decisive insights towards an explanation of individual-level adherence to SGN.

In Section two I summarize the concepts that make Kant's moral psychology relevant to the cases discussed here. I begin by approaching the concepts of necessity and necessitation to present Kant's view concerning the characteristic features of human moral experience. I discuss the concepts of laws or norms to endorse the Korsgaardian thesis that agents' practical identities are sources of moral obligations. Section 03 explores SGN's normative aspect. I contend in this section that it is precisely this aspect that places such narratives on an essentially moral register, and ultimately that looking at the decision-making on green narratives from a moral perspective brings to the surface the problems in endorsing such narratives. I then approach virtue-signaling and self-affirmation to argue that such psychological phenomena find equivalence in Kant's

thought, being regarded by Kant as blameworthy insofar as they hinder the proper functioning of practical reason.

The debate around SGNs can undoubtedly be enriched by a Kantian-inspired contribution. Kant's concept of *maxim* advances the notion of motivation and importantly connects with the idea that our moral life is essentially normative, as it involves the idea of principle; these principles are in play when we make decisions on moral matters. They are enacted by what, in her theory of self-constitution, Korsgaard calls *moral* and *practical identities.*

Contra amoralist arguments[1] I argue that decisions to endorse SGN are typically based on moral grounds, and most importantly that understanding the normative and self-regarding aspects behind such decisions is critical if we want our actions to have moral worth. In this sense Korsgaard's theory is especially appealing, as it provides us with understanding about how practical identities play a role in an agent's reasoning with regard to green narratives, as well as concerning the tensions stemming from our practical identities once we perform actions that endorse narratives we acknowledge as deceptive.

SECTION 1. ACCOUNTING FOR MORAL BEHAVIOR VIA DESCRIPTIVE THEORIES

The increasing interest in awareness of different sectors of society for issues related to sustainability (Zhang et al., 2008; Follows & Jobber, 2000; Shepherd et al., 2005) combined with the growth of so-called ethical consumption (De Pelsmacker et al., 2005) has made "corporate greening" a profitable marketing strategy. This new trend, which has encouraged numerous companies to go green, has at the same time intensified companies' needs in making public their attitudes on environmental issues. In this context and aiming

[1] Claims that moral reasons offer no rationale for acting. A notable example of claiming that phenomena such as those discussed here should not be analyzed from an ethics framework, see Moeller, H. G., 2009. Although Moeller does not claim to be an amoralist himself, he does attack the main idea behind my argument, i.e., that ethics is a possible or desirable way to address contemporary quandaries.

to convince the audience about the commitment and positive environmental impact of a company, goods, or services, green narratives emerge. Many of these narratives came out as merely symbolic or performative, meaning that they fail to comply with what is publicized[2].

In this paper, I refer to the concept of *symbolic green narratives*[3] (SGN, for short) as comprising discourses or communication mechanisms that seek to create or reinforce misleading beliefs in their audiences regarding the environmental commitment of those who produce them. The production of SGN includes not only the manufacture of certain goods, services, or corporate guidelines, but also the discursive manner in which these products are made public. Thus, the scope of what I refer in this paper to as green narratives may include texts, images, packaging, advertising assets, audiovisual media, and other discursive ways of conveying meaning through language.

In the last decades one may witness the emergence of critical stances towards such narratives[4] as well as a growing body of studies addressing this phenomenon more directly. Among the aims of the debates on SGN, the need to list the factors that explain the individual-level adherence to such narratives is typically stressed and commonly carried out by applying certain theories to the scope of the phenomenon under analysis. Two theoretical strategies are worth highlighting.

In seeking clarity about the processes involved in the effectiveness of SGN, Nyilasy et al. (2012) draw on a conjunction between attribution theory and attitude change theory. The former theory, borrowed from general psychology, holds that agents tend to attribute causes for observed phenomena. However, agent engagement with a company's green

[2] Contemporary literature has generally approached these narratives under the label of *greenwashing*. See also note 04 below.

[3] In line with the terminology coined by Bowen, F. (2014), I will accordingly employ the concept of symbolic green narratives to delineate all kinds of intentionally produced discourse seeking to mislead the public about practical commitment or involvement toward environmentally related issues, making it appear that this company, good, or service appears more sustainable than it actually is.

[4] The increasing pervasiveness of misleading environmental communication is reviewed by Lyon, T. P & Montgomery, A. W. (2015). For a critical approach to the phenomenon, see Maxwell, R. & Miller, T. (2017).

advertisement is not immediate; this allows that stimuli caused by green advertising result in a change in the consumer's attitude towards that company. Effectively, this means that when it comes to SGN, the public suspects at the outset that the corporation's pledge is not authentic. This points to the fact that while companies manufacturing SGNs expect positive brand attitudes from their audience, that audience's perception of the merely symbolic nature of these narratives often results in rather negative attributions.

In this sense, one may claim that the explanatory potential about SGN' effectiveness lies in the individual's perception of green marketing, which in turn is mediated by attribution mechanisms and attitude change or formation. As such, once their misleading nature is perceived, these narratives tend to produce a negative brand perception, which brings damage to the firm that adopts them, indicating that "some firms would be better off staying silent" (Nyilasy et al., 2012, p. 121).

A further theory used when it comes to determining the factors responsible for the effectiveness of green narratives at the individual level is the theory of competitive altruism. This theory focuses on a typically pervasive behavior, that of virtue signaling, which has more recently been explored in the philosophical literature under the name moral grandstanding[5].

The theory of competitive altruism is used by Mitchell & Ramey (2011) to outline a hypothesis about consumer adherence to green products, which in turn would arise due to the social sensitivity that individuals generally hold towards their peers. Mitchell & Ramey stress that social and anthropological factors underlie individuals' adherence to green narratives in general, many of which are merely symbolic. Therefore, when certain companies produce green narratives but lack a genuine alignment between their environmental commitment and their policies, regulations, or practices, they rely on the support of such mechanisms to meet their market goals.

[5] The concept of moral grandstanding is defined by Tosi & Warmke (2016, p. 200) as involving the desire "that other people recognize her as morally respectable." In this regard, see also Blackford, R., 2021.

Thus, according to Mitchell & Ramey's application of theory of competitive altruism, the elements accounting for SGN's effectiveness in being endorsed by their audiences are essentially anthropological and social.

While I grant Nyilasi et al that one of the elements explaining adherence to SGN is the individual's perception concerning the falsity of these narratives, I argue that SGN's success may falsely suggest that the role played by the individuals is restricted to the perception that they are potential victims of deception, leading them, as a result, to change their attitude accordingly. Put another way, overemphasizing the perception of falsity might lead one to believe that when perceived, the merely symbolic nature of a narrative *necessarily* produces in the individual a negative brand attitude, prompting the agent to boycott the green deceiver.

A more robust explanation of public adherence to SGN still seems to be needed, in particular, an explanation that accounts for cases that paradoxically result in *positive actions* of engagement *despite* the perception of their symbolic nature[6]. Although not explicitly stated, I believe that explaining how such narratives succeed is one of the goals of Mitchell & Ramey's approach resulting in the application of the theory of competitive altruism to greenwashing cases. This theory would explain why agents tend to undertake actions that display their positive engagement with narratives they perceive as deceptive. In particular, the performance of these actions is directly parasitic on the fact that "individuals selfishly tend to be perceived as altruistic because it elevates one's status, which in turn affords one benefits associated with the higher status" (Mittchel & Ramey, 2011, p. 42). However, while competitive altruism theory contains a thesis that seems to account for why agents endorse SGN *beyond the perception of intended deception*, the way it is framed leaves aside a critical aspect of agents' decision making when it comes to such issues. Namely, it overlooks the fact that such deeds may be pursued not just out of

[6] Put in terms of greenwashing: when perceived, greenwashing results in negative attitudes towards those who produce it (Zhang et al., 2018). Thus, for cases where even if perceived, greenwashing maintains its success in deception, it is still necessary that the factors responsible for this success be determined.

a need for external recognition, but also out of psychological coping mechanisms of self-affirmation.

In what follows, I take a similar path to the aforementioned scholars and offer a hypothesis about individual-level adherence to SGN. My hypothesis complements the theoretical lines I have just discussed in that it advances the following premises:

(1) when it comes to green consumption, individuals' decision about engaging with SGN occurs within an ethical framework;

(2) once in the ethical realm, actions stemming from such engagement (e.g. the consumption of symbolically green products or services, or the reproduction or validation of SGN) may be deemed as moral, immoral, or prudential, as they concern the agent's autonomous deliberation; and

(3) in light of the perceived symbolic or deceptive aspect of these narratives, individuals' adherence is significantly influenced by first-person psychological, in addition to (third-person) anthropological and social aspects.

My goal is to draw close attention to psychological and morally informed aspects of an agent's decisions, in particular, how the practical identities of these agents shape their reasoning with regard to green narratives. A first step towards this will consist in understanding the main traces of human moral deliberation, notably how moral judgments are guided by practical principles.

SECTION 2. KANT'S MORAL PSYCHOLOGY

Imagine a consumer who, out of environmental concerns, decides to stop buying single-use plastic products. Despite her urge to get those products, she suppresses her desire so that her action is consistent with her best judgment. For Kant, such an agent is subject to the power of elements that *necessitate*, i.e., constrain her will, in such a way that she feels compelled to carry out certain actions.

One might, however, wonder why someone *feels bound* to comply with it. Addressing this concern amounts to exploring the nature of what has been discussed under the name of *moral obligation*, which in turn permeates agents' reasoning when they deliberate about moral issues.

Drawing on the idea that as human beings, we have the capacity to act from the representation of laws, in what follows I elaborate on how Kant's concept of obligation captures this distinguishing feature of human nature.

2.1. NECESSITY, NECESSITATION AND MORAL OBLIGATION

In *Groundwork for the Metaphysics of Morals* Kant states that whereas "everything in nature works according to laws" we, as rational natural beings, have "the capacity to act *according to the representation* of laws, i.e., according to principles, or a *will*." (GMS 4: 412).

With this, Kant wants to draw attention to the difference between, on the one hand, the way nature works and, on the other, the way beings that are not only natural, but also rational, act. A lion who kills a gazelle out of hunger is directly determined by sensible impulses, that is, natural ones. Their "choice is not free, but necessitated by incentives and stimuli" (VE: 344). Human moral agency, on the other hand, is not *purely* a matter of a natural law because we are not only sensible, but also rational beings. Thus, our moral actions do not follow immediately from sensible stimuli, meaning that between our desire to eat the remains of animal corpses and the performance of such an action, there is room for deliberation.

Yet, assuming that when it comes to human moral agency there is a space for deliberation involves acknowledging that such deliberation is typically determined by factors that stem from our nature as *sensible* and *rational* human beings. Reflecting the duality of human nature, Kant calls these factors *laws of freedom*, i.e. laws that concern free human action, constraining it. Those are, as the concept itself emphasizes, free actions that are

also marked by laws; laws that nevertheless do not imply natural necessity (such as in the case of natural laws) but rather necessitation [*Nötigung*].

Kant makes the relationship between *free action* and *necessitation* explicit in the opening sentence of the Introduction to his *Doctrine of Virtu*e, where he maintains that "the very concept of duty is already the concept of necessitation (constraint) of a free choice through the law" (MS, 6: 380). To see what this means, we turn to the elucidated passage in the *Groundwork for the Metaphysics of Morals* where Kant defines *willing* as our ability to act in accordance with the representation of laws. This equivalency[7] means that, for Kant, human will is oriented to change – and, in some sense, *to cause* - states of affairs in the world. In turn, being able to act from laws means that one's actions are not mere movements arising from stimuli and sensible impulses, but rather byproducts of intentional forces, which point to precisely what must be taken into account if reason is to acquire a practical interest. As Kant points out,

> [t]he human power of choice is indeed an *arbitrium sensitivum*, e yet not *brutum* but *liberum* because sensibility does not render its action necessary, but in the human being there is a faculty of determining oneself from oneself, independently of necessitation by sensible impulses. (KrV A534 / B562)

A lion (which has *arbitrium brutum*) does not choose to hunt its prey; it is instead immediately responsive to what is necessarily determined by its sensibility. Unlike a lion, human action is not *directly* determined by sensibility. Conversely, when morally worthy, human action can be carried out on purely rational principles.

To see why this is so, take Kant's concept of incentive [*Triebfeder*]. For Kant, human action always involves incentives, because as human beings, we need incentives in order to be motivated to act. An incentive is, as Korsgaard (1996, p. 242) clearly puts it, a kind of "first-order impulse (...) to the performance of an action". But to count as a motive for a certain action, incentives still need *principles*. According to Kant's moral psychology, the principle behind our choices may ultimately be either rational, i.e., *duty* or material,

[7] Which ultimately refers to the equivalence between will and practical reason. See *GMS* 4: 412.

i.e., *self-love*. Put another way, we say in the former case that an action is performed for the sake of duty, and in the latter, out of inclination [*Neigung*].

Such a view with regard to the principles of choice has two notable aspects. First, as said, that at bottom, all moral action is ultimately grounded on what is required by *either morality* (rational principle) *or self-interest* (material principle).

Only in the former case such an action bears moral worth. For example, one may resist the urge to buy single-use plastic products out of duty *or* out of inclination. If the agent's concern for the environment is what determines the action, her action was done from the motive of duty, thus bearing moral worth. In other words, her action has moral worth because the rational principle, i.e., her observance of what morality requires, is what caused her to decide to act in a way that curbs her inclination towards single-use plastic products. On the other hand, if one's action was merely performative, with an eye to, for example, making the agent appear to be more virtuous than she really is, then her action (although in accordance with duty) has no moral worth.

But what is wrong with an action being done out of inclinations? What if the effects of my actions are good, even though their determining motive is, for instance, my sympathy towards green issues? For Kant, for an action to have moral worth, it cannot have been performed out of material principles, because those are merely conditioned. It is conditioned, i.e., contingent, that I am an environmentally sympathetic person, as I could of course be a person who does not care about green issues[8]. Therefore, this cannot be the determining reason why I perform my action, if it is to have moral worth. Analogously, the moral worth of an action is unrelated to "the *effect* that is expected from [that action], nor therefore in any principle of action that needs to borrow its motivating ground from this expected effect" (GMS 4: 401) since such effects are similarly conditioned, meaning that they "could also have been brought about by other causes". The moral worth of an

[8] Kant's illustrious example of the philanthropist (*GMS* 4: 398) is widely used in the literature when it comes to these distinctions. It deals with a person with a sympathetic temperament - whose compassionate action finds intimate pleasure in performing benevolent actions.

action is closely tied, for Kant, to what is distinctive of a rational will, namely the ability to act according to the representation of laws and therefore from the sake of duty.

A second notable aspect with regard to the principles of choice is that principles are able to determine the will (and therefore one's deeds) only insofar as they are incorporated into actions as a result of one's free choice (REL 6: 24).

Thus, while we may often find our actions strongly driven by sensible impulses such as passions and other pathological emotions[9], we can still act as morality requires. Yet, whatever incentives present to us, they only become the determining motive for the action after being incorporated in that action as a result of our self-determination. This means that an agent must render such an inclination as a part of the *principle* one gives to oneself. Thus, no matter how much one's action is naturally bound to sensible impulses, one must have incorporated those incentives into what one regards as the principle for which one acts. Such a self-imposed principle Kant calls a *maxim*. Agents that conform their actions to, say, environmental preservation, hold this practical rule as their maxim. This points to the fact that *human moral actions are principled*, meaning, they bear maxims[10].

Yet, maxims allow us to assess our actions. As was discussed earlier, as agents we act for reasons, that is, for inclinations we represent as sufficient motives for our actions, meaning that we *incorporate* certain inclinations into the maxims of our actions. In effect, as agents, we experience and act under the idea of freedom to the extent that freedom is imbued within the idea of agency. Not to consider oneself an agent amounts to not interacting with the world at all. Given that good actions, that is, actions with moral

[9] Cases of *practical akrasia* (weakness of the will i.e., the performance of deeds against our best judgment), highlight the interference our sensible impulses have on our decisions about what we have most reason to do.

[10] The concept of a maxim is not widely discussed by Kant, for there was at Kant's time a certain consensus about its meaning. For a broader understanding regarding the concept of maxims, in particular *practical maxims*, see Kitcher, P., 2003. She addresses, among other relevant aspects, the apparent contradiction that arises in Kant's writings as he considers maxims as bearers of moral worth, other accounts on this concept, as well as the understanding of traditions preceding his ethics.

worth, are those performed from duty, ensuring that interest in morality, i.e, duty, is the determining motive for that action is tantamount to confirming that we are on the right path to performing a good deed. Hence, the maxims of our actions allow us to evaluate whether or not our actions have moral worth. Stated another way, the maxims of our actions make it possible for agents to assess how they act in the world.

Of course, it is not the case that we should only act from maxims that allow moral assessment. It is true that a number of actions we perform do not have a moral dimension, and Kant himself calls attention to over-moralizing issues that do not merit moral consideration[11]. However, once we genuinely acknowledge a topic as morally charged, we must seek to act out of duty, despite the strong tendency to act according to our inclinations.

Now, back to the concepts of law and necessitation. As such, our experience of free agency is essentially connected with the idea of laws. Laws give shape to this relationship. As Korsgaard (2009) emphasizes, they shape our daily actions. Norms defining one's "*do's and dont's*" within a social democracy, the parental demands regarding their children's education, and other sets of norms that we follow on a daily basis, define our social life by means of setting boundaries on our total freedom as individuals[12].

However, sometimes,

> we find ourselves doing what we think we ought to do, in the teeth of our reluctance, and even though nothing obvious forces us to do it. We toil out to vote in unpleasant weather, telephone relatives to whom we would prefer not to speak, attend suffocatingly boring meetings at work, and do all sorts of irksome things at the behest of our families and friends. (Korsgaard, 2009, pp. 2 - 3)

[11] Cf. *MS:* 409. Kant calls "fantastically virtuous" one who is overly concerned with non-moral details, such as "whether I eat meat or fish", an example that is however, easily moral, since it involves taking the lives of sentient beings who have an interest in their existence.

[12] In this sense, such norms refer to more abstract ideas and goals, that is, values that guide human behavior and the way individuals assess it.

The examples outlined by Korsgaard suggest that many of our actions are carried out as the outcome of laws or commitments we set for ourselves. They are, in this sense, *internal*, and represent what an agent sets as *a norm or law for oneself*[13].

Thus, saying that, for instance "one *ought* to x" illustrates a principle. This principle involves that our action can go one certain way, but not another, meaning that our action is constrained, i.e., necessitated. We identify the outcome of these constraints with what we refer to as *duties*. In this sense, the very concept of duty entails a constraint to our will. That is what obligation is based on: the concept of duty implies from the outset the idea that a will that suffers necessitation in its freedom is under obligation imposed by reason.

By extension, laws of freedom express duties that necessitate our will to perform actions that we would not have pursued if those norms were not posed. For example, when I refrain from getting a product I would like to have because it comes packaged in single-use plastic, there must be something compelling me to act in this way, since if it depended solely on my desire, I certainly would not do so.

Kant's account of obligation captures these features of our moral psychology. According to him, ethical norms are likely to be represented by us as commands or *imperatives*, which, in turn, necessitate our will. And here is where the distinguishing factor of the typically human moral experience lies: these norms count for sensible and rational beings like us. Only human beings are capable of acting from the representation of laws, i.e., of understanding and acting on imperatives. Such imperatives place human beings under obligation, binding them by virtue of their (rational but at the same time sensible) nature. If we were merely rational beings, moral norms would not apply since our will would already conform to what is morally required. Conversely, when it comes to laws of nature, the idea of obligation does not seem to apply. One is not likely to claim that a lion

[13] Human freedom can also be hindered by *external*, i.e., juridical or legal laws. Both sorts of duties (internal or external) are capable of imposing norms that necessitate, i.e., create restrictions or constraints on human will.

feels obligated to hunt its prey, for what determines the occurrence of such a natural phenomenon is a cause-effect necessary relation. It is thus reasonable to claim that the idea of obligation amounts to a distinctive aspect of our moral experience. There is a sense in which we say that we are obligated to treat our fellow human beings respectfully, to look after people we love, to provide help when we are asked to do so, and so forth.

These distinctions shed light on the idea of moral obligation or, put differently, on what determines the sense of obligation we experience when dealing with particular circumstances. Once in the moral realm, we make decisions that not only involve laws, but depend on our distinctive way of apprehending and representing such norms.

In the next section I will argue that when it comes to SGN, the way we experience these obligations, namely a typically moral way, is a prominent and necessary factor in understanding what lies behind our endorsement of such narratives.

SECTION 3. EXPERIENCE AND MORAL OBLIGATION ON SYMBOLIC GREEN NARRATIVES

In broad terms, the aim of marketing strategies amounts to generating and maintaining competitive advantages through tactics geared towards satisfying the needs and wants of a target audience (Obermiller et al., 2008). Concurrently, manufacturing green narratives can be seen as one of such tactics, for they are designed to appeal to consumers' sense of obligation as a path to generating engagement.

The fact that consumers' sense of obligation is central to such narratives' success points to the ethical nature of the decisions and actions arising from them. However, from the fact that these narratives are *designed* to prompt a moral appeal, it does not follow that the decisions and actions concerning them have moral value from a Kantian perspective. This is because the correct motivation is crucial for moral value; the factors involved in the production of green strategies are not inherently associated with the decisions and actions that may eventually result from these narratives. In this sense, the presence of an element

intrinsic to the decision-making process would be necessary to ensure that such actions and decisions are within the moral realm.

My aim in the next section is to introduce this element. I elaborate on the idea that decision-making about green narratives must be faced as a decision that encompasses a moral aspect in virtue of its normative character. I will show that once we focus on how market strategies are designed and performed by means of targeting a certain type of consumer, the essentially normative character of these decisions and actions emerges, a fact that emphasizes the sense by which *the way* these narratives are produced plays a heuristic role. I contend that we may conclude that once perceived, these narratives connect with consumers' practical identities, producing moral obligations. In the last part of the section, I rewind to the idea of moral worth, and provide reasons why a comprehensive analysis of the phenomenon must take into account a first-person perspective that embodies explanations about psychological aspects that may act as impediments in assessing our maxims. Finally, I discuss why failing to carry out such appraisals may raise problems.

3.1. THE ETHICS OF SYMBOLIC GREEN NARRATIVES

The emergence of ethical consumption and its prominence, particularly in recent decades, has led companies to pursue green marketing strategies. In general, market strategies are devised based on data concerning the attitude and consumption patterns of a certain group of consumers; in the case of green strategies, the targets are green or ethical consumers.

While not coextensive,[14] ethical and green consumption are often addressed simultaneously. Both refer to the behavior of "people who are influenced by environmental or ethical considerations when choosing products and services" (Crowe &

[14] See Sho & Krasser, 2001.

Simon, 2000, p. 04)[15]. It is thus possible to make the general claim that ethical or green agents are those whose decision-making incorporates environmental and social considerations. This means that the actions resulting from these agents' decision-making processes are determined by certain considerations, and that in most cases they would be different if their agents had not taken such considerations into account. For example, ethical consumers do pay more for products or services that are explicitly consistent (as an instance, via labeling[16]) with their moral commitments (Hainmueller, et al., 2015; Hertel et al., 2009); in the absence of moral commitments, the same consumers would arguably have responded differently towards ethical branding. In other words, green agents' environmental and social concerns are a defining component of their actions and decisions concerning green narratives.

Now, in the environmental and social marketing literature the fact that these concerns may in turn be defined in normative terms remains underexplored, as the literature on those fields seeks to be descriptive. Thus, the claim that green consumers make their decisions out of social and environmental concerns is, now in moral terms, equivalent to the idea that the considerations green agents incorporate when it comes to social and environmental issues *necessitates* their will, meaning that their decisions are constrained by norms or laws they set for themselves. Thus, the stance of deciding whether to engage with such narratives, added to the necessitation generated by the attitude towards green or ethical issues, pushes the phenomenon to the ethical sphere.

To achieve a better understanding of this normative element, it may be helpful to, once again, turn to some ideas concerning our typical moral experience, now through Korsgaard's theory of practical identity.

---

[15] Crowe & Simon (2000, p. 04) emphasize that the term 'ethical' covers "matters of conscience such as animal welfare and fair trade, social aspects such as labor standards, as well as more self-interested health concerns behind the growth of organic food sales."

[16] Empirical evidence suggests a positive correlation of ecolabelling on consumer purchasing behavior. Teisl et al (2002) highlight this correlation based upon evidence of the impact of the introduction of the 'dolphin-safe label' in tuna retail sales. More recently, Dhir et al (2021) found that labeling satisfaction is one of the main drivers of green consumption in the context of apparel.

As said, Korsgaard argues that the necessitation we experience in cases such as decision-making on green issues results from duties we set for ourselves. Her agential identity theory draws on Kant's moral psychology to claim that setting duties for oneself is closely connected with one's practical identity[17]. According to Korsgaard (1996, p. 101), the conception of one's practical identity refers to "a description under which you value yourself and find your life worth living and your actions worth undertaking". It involves an agent endorsing the set of obligations attached to a certain way of identifying oneself (Korsgaard, 2009, pp. 22 - 25).

While Kosrgaard concentrates on demonstrating how agency and self-constitution are co-extensive, we do not need to endorse her argument to see that as a matter of fact, our identities are often the sources of obligation, by means of making claims on our actions[18]. One whose practical identity comprises the conception of a benevolent person, for example, does not perform actions that involve taking advantage of other people. Similarly, a person who portrays herself as an environmentalist, that is, someone whose environmental protection is incorporated into her practical identity, must refrain from contributing to environmental degradation.[19]

Our practical identities are, for Korsgaard, merely contingent. We acquire certain practical identities because we are situated in certain places, because we play certain social roles, or because we experience certain situations. They amount to the way *we regard ourselves* and, in many cases, result in obligations we deem as unconditional, that

[17] More specifically, she expands Kant's theory by arguing that practical identities determine moral agents' choices just as, in Kant's view, moral principles do.

[18] Korsgaard's notion of practical identity holds a prominently social dimension, since, as she summarizes (2009, p. 20) it includes "roles and relationships, citizenship, membership in ethnic groups, causes, vocations, professions and offices." It is, in this sense, consistent with empirical evidence about the salience of these social roles and interactions in our moral judgements. In this respect, see Hamilton & Sanders, 1981; Kaspar et al., 2016; and Willemsen et al., 2018.

[19] Studies on environmental self-identity (Van der Werff et al., 2013) support this idea by showing that there is a positive correlation between environmental self-identity and pro-environmental actions.

is, obligations whose violation makes us feel as if we are losing our integrity (1996, p. 102).

There is a form of identity upholding our practical and contingent identities viz., the moral identity. It is nevertheless necessary and therefore, once assumed, a normative source of moral obligations that are likewise necessary. To have a moral identity is for Korsgaard (1996, p. 121, my emphasis) “to value yourself *just as a human being*".[20]

Now, our relationship to our practical identities is, according to Korsgaard, a construct. For her, "every time you decide to act in a way that conforms to your practical identity, you count as re-endorsing that form of identity and making it your own" (2011, p. 09). As such, we can always give up our practical identities and we do so when, for example, we decide to overlook the obligations they generate. However,

> so long as you remain committed to a role, and yet fail to meet the obligations it generates, you fail yourself as a human being, as well as failing in that role. And if you fail in all of your roles - if you live at random, without integrity or principle, then you will lose your grip on yourself as one who has any reason to live and to act at all. (KORSGAARD, 1996, p. 121)

This joint failure occurs because of the relationship between practical and moral identities. Our practical identities acquire normative force because of the requirements of our humanity, that is, to the extent that assuming certain practical identities is part of our constitution as beings who value themselves as human beings.

[20] To bear a moral identity amounts to "valuing humanity in your own person rationally" (*ibidem*) which in turn also implies valuing humanity in the other person's rationality.

From this, it is rendered explicit that normativity makes up our moral experience. As rational yet sensitive human beings, our will is free to the extent that our actions do not necessarily result from our inclinations. A person who desires to eat meat but regards herself as an environmentalist may wish to engage with SGN, purchasing and consuming, say, "fair meat"; as her will is necessitated by the obligation arising from her practical identity, the possibility that her deed results from something other than a necessary cause of her desire to eat meat (which happens in the epitomized case of the lion that kills the gazelle) opens up.

The contrast between necessity and necessitation (or more accurately, this fact about our moral nature) is precisely what is at stake with regard to decision-making overall, and especially, with decisions within the moral realm. Chief among these actions are those that result from agents' decisions whether to engage with green narratives that are perceived as symbolic. Once these agents perceive the symbolic character of such narratives, they often ask themselves "What should I do now?" or "What is the best deed in such a scenario?", which in addition to emphasizing that their will is undergoing necessitation, points to the fact that their decision lies in the moral realm.

From what has been discussed, it is possible to evince that, in the context of green narratives, the individual decision process on whether to engage with such narratives is ethical par excellence, as the individuals' sense of obligation is triggered by virtue of the normative nature of the issues that emerge from their practical identities.

Additionally, it is worth noting that understanding how marketing strategies that focus on these individuals assists us towards *clarifying* that the decision-making process regarding SGN has an ethical character. As mentioned, green strategies are designed from investigation into the behavioral and attitudinal profile of the consumers to be reached. That granted, take as illustration the increasingly adopted strategy of labeling products as "fair", which reflects the market demand known today as "fair trade", a form of ethical

consumption associated with economic and environmental justice[21], as well as animal welfare (Annunziata, 2011). Over the last few decades this terminology has also been used to stress that a product is fair by virtue of one or more (but seldom all) aspects of its supply chain. As an individual who regards the concept of "fair" among the descriptions under which one values oneself in the sense outlined, an agent may feel compelled to engage in a particular way when confronted with SGN. One might for instance decide to consume "fair meat" or, once aware of its merely symbolic character, to boycott this product or brand. The use of an ethical concept such as "fairness" as a green marketing strategy allows us to see that what drives such strategies is precisely the fact that their effectiveness is inscribed in the ethical field. This is the reason why *the way* SGN are produced is relevant, namely, to point towards helping us to better grasp such narratives' normative aspect.

However, SGN manufactured as targeting green/ethical consumers do not always succeed in reaching that audience. That leaves room for cases of collateral consumers, i.e. individuals who end up engaging with such narratives arbitrarily or for essentially contingent reasons, e.g. ignorance or lack of interest. One could argue that the Kantian moral theorizing I bring up here does not succeed in explaining cases of collateral consumers, because for cases like these, it is hard to see how moral motivation could be involved. However, Kantian theory does also allow us to assess collateral consumers' endorsement deeds. Especially, it allows us to see that, *grosso modo*, collateral consumers will not act on these considerations *because* the essential component of green narratives,

[21] Overall, fair trade initiatives are committed to challenging "the ecologically and socially exploitative relations" (Schreck, 2002, p. 13) characteristic of industrial production systems, in order to "provide consumers with a guarantee that products are grown by disadvantaged producers *under healthy social and environmental conditions*" (*ibidem*).

i.e., the environmental one, does not create obligation in them.[22] They may endorse SGN *despite* the fact that those narratives are green.

The fact that moral motivation results from obligation also helps us to understand the apathy of such agents toward green issues, as for when green narratives do generate obligation, consumers tend to feel motivated to perform actions incorporating ethical/green considerations.

For example, I may buy a recycled packaged product and decide not to share this product on social media, once I have learned that it is greenwashed. In the first moment, I endorse, i.e., positively engage with a narrative that portrays a certain product as green. In the second, upon learning that the product is greenwashed, I decide not to platform it, therefore engaging negatively. In both cases, my sense of obligation - in particular, moral obligation with regard to an ethical issue - motivates me to pursue certain courses of action.

Agents' personal identities certainly generate moral obligations, duties as well as inclinations. However, endorsing green narratives must ultimately result from one's own choice to take these obligations as a source of reason to act upon. Thus, following the example, as a person who regards oneself as an environmentalist, one may acknowledge one's duty not to platform a product whose greenwashing is perceived. However, the relationship between representing something as my duty and my actual deed is not a relation of necessity. This means that despite the motivation this duty generates on me, I can still deliberate and act differently, for example by sharing the greenwashed product on social media.

[22] In terms of marketing strategies, that would be analogous to targeting a shampoo campaign at bald consumers. The theory I outline here maps the behavior of these people to the same extent that the theory of competitive altruism maps the behavior of people who truly don't care about being perceived as non-virtuous, non-altruistic or non-empathic.

This denotes the importance of the agent's intentionality in *choosing to engage* with certain narratives, which in turn further reinforces the line that divides such behavior (disinterested or apathetic) from the ethical one.

To sum up, while the fact that narratives designed to work in an ethical framework produce decisions and actions of (positive or negative) engagement inscribed precisely in this framework is not surprising, it can still be misleading. Accepting that x is ethical because it is designed to prompt ethical responses does not sufficiently account for the fact that these responses are necessarily ethical. Inscribing actions of endorsement or engagement in the ethical field further depends on elements *internal* to those actions. When we make decisions based on the specific practical identity's conception targeted by green narratives, we bring out this internal element: the normative character our experience acquires when faced with such narratives. Engagement with green issues depends on the moral obligation generated by such narratives; only then are they able to motivate us to pursue actions that incorporate green considerations. Therefore, although the external element does not guarantee that these narratives take place within an ethical framework, it is (through the idea of target audience and thus of practical identity) a key heuristic piece to elucidate the ethical nature of decision-making regarding SGN.

3.2. MORAL DENIAL AND SELF-DECEPTION

Actions stemming from agents' sense of obligation are actions within the moral field. But that is not to say that they are actions with moral worth in the (Kantian) sense sought above[23].

When an agent engages with green narratives, her deed bears moral worth only insofar as it premises the compliance with and respect for the obligation posed. As discussed, an action has moral worth when performed *out of duty*. In contrast, actions that merely conform to duty or go against it bear no moral worth; for an action to have moral worth,

[23] See section 2.2.

the agent's will must be necessitated, restricting sensible impulses, and rendering duty the determining motive of the action.

To elucidate the idea of moral worth, take as illustration the behavior of an agent faced with a SGN, for example, a product labeled as "fair meat". As an ethical consumer, this agent tends to ask herself "what should I do in this situation?", and despite the perception of the merely symbolic aspect, she may decide to endorse it. In light of her awareness that its production inevitably violates ethical and environmental principles, she may decide to purchase the product despite uncertainty as to the validity of attributing "fair" to "meat". As an individual whose conception of practical identity incorporates environmental considerations, she violates the duties and obligations engendered by her practical identity. In Kantian terms, this amounts to saying that she violates moral requirements. Thus, despite awareness of these requirements, that agent allows her will to be determined by *something other than duty*. Her motivation seems to be rooted in the fact that she desires the benefits associated with that product[24] in such a way that the pursuit of those perceived benefits drives her behavior. Consequently, by violating moral requirements, her action lacks moral worth. Therefore, when her desire to consume "fair meat" overrides what she acknowledges as her duty, or put another way, the moral demands set upon her practical identity, then she acts out of self-love[25].

As such, overall, actions performed out of self-love and despite the agent's acknowledgement that these actions violate their moral requirements, at the same time infringe upon these agents' practical identities.

[24] Benefits typically perceived by meat eaters are, for example, aesthetic - related to the taste of the meat (Piazza et al., 2015), social - related to the social impact within gendered food consumption ideologies (Adams, 2018; 2007), socio-environmental - aiming at achieving acknowledgement of ecological virtues by her peers, etc.

[25] It is to be noted that one does not have to fully endorse a course of action in order to pursue it. Examples are cases of akratic actions, i. e. those one pursues despite the awareness that they violate one's moral requirements.

Now, violating one's own conception of practical identity in turn yields a threat to one's moral integrity. Take again the case of green or ethical consumers by imagining that a green consumer regularly breaches the requirements and duties engendered by her identity. As a result of a reflexive process (Powers, 1973), and because such violations are associated with negative self-attributions (Giner-Sorolla, 2001; Kivetz & Zheng, 2006), that individual tends to challenge the validity of her subsumption to such a practical identity's conception. Since that identity is an aspect of one's self (insofar as it represents a description whose value is initially attributed by the individual oneself), its violation simultaneously amounts to a threat to one's own integrity[26].

This idea is consistent with contemporary theories of personal identity, which define and typify the conflicts arising from the discrepancy between actions and thoughts, in particular between agents' behavior and what they think about their own identity (Carver & Scheier, 1981; Stets & Burke, 2000). Following the framework, I am outlining here, this is to say that these conflicts emerge from the lack of agreement between actions and decision-making on the one hand and moral requirements on the other.

Now, we can grant that in the context of green symbolic narratives, one may resolve this conflict by means of simply *abandoning* what generates the moral obligation, i.e. a certain conception of one's practical identity. Examples include cases in which individuals forego a plant-based diet on the grounds that the costs associated with that value are too high.

Yet a different path to address this conflict, or rather to cope with the resulting dissonance, depends on the agent's choice to *preserve* the conception of practical identity that leads to the dissonance. Should the agent preserve such a conception while respecting the moral demands it entails, the dissonance dissolves. Conversely, it becomes problematic once that agent chooses to preserve a certain conception of her practical

[26] Cf. Korsgaard, 1996, pp. 100 -102.

identity *despite* the competing demands it engenders[27]. Upholding a conception of a practical identity amounts to meeting the obligations that such an identity prompts; doing so while engaging in actions that *violate* that very conception is akin to having the cake and eating it too.

It is within this context of discomfort produced by the lack of consistency between one's actions and obligations that psychological mechanisms of self-preservation emerge. Strategies such as rationalization, wishful thinking and self-deception, are coping mechanisms that aim to preserve the agent's moral unity by reducing the cognitive dissonance entailed by the inconsistency between the awareness of nature and the performance of immoral deeds.

However, despite the typified advantages,[28] such strategies hold potential harm. They can play a critical role in distorting one's moral reality, as they can make immoral deeds look like permissible deeds that would otherwise be disallowed.

These phenomena are discussed by Kant under the label of rationalization or self-deception. Moral rationalization or self-deception[29] is, according to Kant, a form of lie (inner lie) that we tell ourselves for exculpatory intent.

---

[27] There are multiple reasons why an individual may want to preserve practical identities whose demands are not wholly in agreement with his or her desires. Because they are understood and assessed within cultural contexts, those reasons are typically associated with the social sphere, but they are correspondingly often related to those individuals' self-perceptions. Kant acknowledges the importance of the social sphere in the context of demonstrating virtue. For example, in *Religion within the Boundaries of Mere Reason* (6: 27), he describes what he calls a predisposition to humanity as a predisposition to "gain worth in the opinion of others," which can become a vice to the extent that the agent comes to desire "superiority for oneself over others."

[28] Such as avoiding pain, stress and anxiety generated by the threat to integrity (Goleman, 1985; Smith, 2004).

[29] I take both terms as coextensive. For a background on the term rationalizing [*Vernünfteln]* in Kant's theory, see Sticker, M. 2021, pp. 08 – 17. For a discussion on the placement of *self-deception* in Kant's work, see Papish, L., 2018, pp. 68 – 69.

For deeds endorsing SGN, we frequently deceive ourselves in order to preserve our moral integrity. There are different ways of engaging in self-deception, and some of them are discussed by Kant throughout his moral writings. For instance, we deceive ourselves once (after having acknowledged a moral imperative) we create exceptions for ourselves with a view to rendering permissible a deed initially disallowed by that imperative (GMS 4:242)[30].

Another instance of self-deception that Kant addresses in his moral theory concerns its epistemic aspect. This form of self-deception emerges once one distracts oneself, i.e., deflects attention from aspects that otherwise would count as morally relevant to the decision whether to engage in a certain action.

Take for instance a green agent who positively engages with a "fair meat" narrative while perceiving its purely symbolic character. Suppose that she motivates her decision on the idea of *preference*, meaning that, for her, this issue is allegedly purely aesthetic such that the decision regarding whether to engage with such a narrative does not require moral scrutiny. Yet, as a green agent, she is aware of the conflict emerging from the deed she is about to perform. It may be that she is provided with information on the impact of meat production upon the rainforest, on the sheer amount of water required to produce 1kg of meat, or data on greenhouse gas emissions related to the livestock industry, and so on. To successfully alleviate discomfort, such information needs to be suppressed. Focusing on the aesthetic aspect in order to take the issue out of the moral arena, ignoring or selectively interpreting information that points to the moral framework are some strategies agents usually employ to ensure this success. What these moral-epistemic strategies have in common is that self-deception seems pervasive in all of them.

[30] Instances of such behavior are individuals who classify themselves as vegetarians but violate the dietary requirements for self-attribution of that identity. For an empirical investigation of the factors leading to such dietary violations, see *Rosenfeld & Tomiyama, 2019*. Through a qualitative study, they address the ways in which meat-eating vegetarians cognitively cope with dietary violations, emphasizing the *post-hoc* justifications following from that behavior.

For Kant, this behavior is blameworthy because self-deception renders the agent who performs it incapable of truthfulness regarding one's assessment of that discrete maxim, as well as of one's moral character. By intentionally manipulating the truth, the agent harms her idea of herself as a truthful person. Ultimately, what is at stake here is the agent's interdiction with regard to the proper use of her practical reason[31]. This is why Kant, in describing self-deception, does so in contrast to external lying, for both violate the principle of truthfulness.[32,33]

Now take the behavior of virtue signaling, a typically others-oriented behavior that has been extensively addressed in the literature.[34] Agents signal virtue once they provide positive moral information on their alleged virtue to people whom they would like to believe that information. Virtue signaling behavior can be genuine, but it can also be merely symbolic. In the latter case, the information an agent provides to one's peers is *misleading*.[35] In this case, it can be construed as an external lie one tells others concerning one's own moral characteristics.

As discussed, competitive altruism theory avails virtue signaling to account for instances of behavior from which actions of engagement with symbolic green narratives are instantiated. Nonetheless, the way competitive altruism theory is typically rendered when it comes to explaining endorsement of narratives such as greenwashing, overlooks an aspect that has recently been shown to be pervasive when it comes to actions within the moral field, which is *self*-signaling.

---

[31] I will get back to this issue later when I discuss the priority of self-deception over external lying.

[32] I discuss the relationship between external lying and self-deception elsewhere (Zanchet, 2022), where I argue for the centrality of the duty of truthfulness within Kant's ethical writings.

[33] This principle is central to Kant's account of duties that an agent has towards oneself because (among other reasons), the agent's moral improvement depends on it.

[34] The practice of virtue signaling is not a practice restricted to individuals. Companies may also engage in virtue signaling (Berthon et al., 2021). As a matter of fact, business models conveying positive signaling stand out due to their potential to improve companies' green credentials (Attah-Boakye et al., 2022).

[35] It is worth noting that here the idea of lie does not concern the negative sense of truth by equivalence, that is, it does not necessarily involve verisimilitude. Instead, a lying proposition is one that fails to carry truthfulness. Accordingly, what is relevant here is that the agent *acknowledges* certain content as misleading.

Parallel to external lying (virtue-signaling), self-signaling can be understood as a lie an agent tells *oneself*, a form of self-deception, or in Kant's terminology, an inner lie. Self-signaling, i.e., the behavior of signaling virtue *to oneself* is, in this sense, analogous to self-deception broadly construed, and, interestingly, to the moral entendre under which Kant's account of the phenomenon is inscribed. Only agents who have some interest (even merely performative) in morality wish to be virtuous.[36] Moreover, it seems that virtue-signaling may at the same time embody self-signaling: by signaling virtue to others, an agent *reinforces* her assessment regarding her own moral traits.

In his account of deception in the *Doctrine of Virtue*, Kant gives priority to self-deception over external lying. For him, external lying necessarily runs through a deception of the agent against oneself. In deceiving the other and thus, in making impossible the proper use of other's moral capacities, the agent violates one's own practical reason. These are reasons why immoral actions in general, and actions embodying deception in particular are regarded by Kant as irrational. Such actions imply that the agent is not making proper use of her practical reason insofar as she fails to enact the reasons produced by the rational process. Her rational process is flawed insofar as she fails to take into account the norms of evidence formation, for instance, by shifting focus away from the moral aspect of the information that constructs her beliefs (Papish, 2018)[37].

Those familiar with Kant's account of lying in a superficial sense characteristically assume that for Kant, lying is inherently wrong because it (1) amounts to a misuse of one's status as an end in itself while at the same time (2) threatens the possibility of making promises. In the latter sense, an agent who promises falsely precludes the

[36] Such a fact is crucial for the emergence of moral self-signaling and it is even more evident when it comes to cases of *virtue* signaling. This is additionally consistent with Kant's views on rational interests. As Sticker, (2021, p. 30) emphasizes, a rational interest that would function as an incentive for moral actions would be the rational interest an agent has to feel rationally justified. This interest, he claims, "is rooted in an agent's acknowledgement of the authority of duty", that is, in morality.

[37] In the third chapter of her book, Papish highlights the aspect of rationalizations that concerns the violation of belief formation norms. Once an agent is unable to contradict what she is already aware of, she may resort to psychological strategies to shift her attention to cognitions that are more attractive to her.

existence of that speech act in a world where falsely promising is a universal maxim. However, this is not the moral problem posed by lying. Lying is morally troublesome because it amounts to the causal manipulation of another's will.[38] By lying, an agent prevents the deceived person from accessing the reasons that ultimately justify her own action, which in turn incorporates into her action (and thus into her practical reasoning) the result of the deceiving agent's maxim. This shared structure between deceiver and deceived is clearly vicious, for it leads the one who receives the lie to generate a maxim lacking a complete justification.

That said, drawing the parallel between virtue-signaling and lying becomes inescapable. In the same way that virtue-signaling can be regarded as a form of external lying, moral self-signaling is a form of self-deception, that is, a lie we tell to ourselves that prevents us from truly accessing and being able to evaluate our practical reasoning.

Insofar as it obscures the exercise of practical reason, self-deception is bound to have negative consequences for the agent who rationalizes away all discomfort caused by the perception of the potential immorality of her actions.

Self-deception can also lead to cognitive losses, such as emotional desensitization (Garrett et al., 2016), and may furthermore decrease the ability of self-deceived agents to successfully pursue informed rational judgements, which can in turn lead to less accurate predictions (Chance et al., 2011). In terms of moral capacities, self-deception is especially pernicious, as it subjects the agent to a distorted view of moral reality. This is problematic because, given the relevance of ethical considerations in consumer decision-making overall (Chowdhury, 2017; Martinez & Jaeger, 2016), self-deception precludes the most critical tool to combat deception promoted by companies making use of green strategies: moral education.

[38] I follow, in this regard, the interpretation outlined by Herman, B., 2022.

Moral education, for Kant, is essential not only for pursuing virtue[39]. It is also key when it comes to forming a moral community of responsible agents. Without the ability to define and pursue ends, to form informed rational judgements, and to assess one's actions and moral character, no moral progress is possible. In addition, the absence of moral knowledge precludes moral education and, as a result, the formation of such a moral community is jeopardized.

Most importantly, self-deception represents a critical risk for the very exercise of morality. This is one of Kant's most central contentions in the *Doctrine of Virtue (MS 6:429 - 431*). Along with external deception, self-deception corrupts agents' ability to use the capacity from which their personality is derived, namely, the ability to properly use practical reason. This happens because both lying and self-deception manipulate one's will, obscuring the way the agent regards oneself and preventing one from acting from duty. Analogously, actions whose maxim is signaling to oneself one's own supposed virtue pose a similar risk.

In the context of ethical agents' decision-making in regard to SGN, self-signaling renders unnecessary any scrutiny that might eventually lead agents to care more than symbolically about morality. Moreover, by signaling their virtue to others these agents flag to themselves an alleged conformity to their practical identities in a process of self-deception that may in turn lead them to fail to enact the reasons why they embraced such identities in the first place.

[39] MS 6: 477.

CONCLUSION

From what has been said above, the reasons why it is important for an agent to be able to evaluate one's maxims emerge. By assessing maxims, agents are able to ascertain the conformity between their actions and the practical identities they have embraced. Moreover, they are able to ascertain whether discrete actions conform to what morality requires, that is, whether they have moral worth or, put another way, whether such actions conform to one's moral identity. In this regard, evaluating maxims is especially important because we cannot guarantee that the practical identities we adopt will necessarily generate morally worthy or nonconflicting obligations. That is to say that given the contingent character of whatever practical identities we adopt, one might conceivably adopt identities such as that of an egoist or a mafioso, which in turn would generate obligations accordingly[40]. In this scenario, the process of evaluating maxims is critical to ultimately bringing agents that hold immoral forms of self-identification back into morality[41].

Of course, self-deception itself threatens our ability to assess our maxims, another reason why it is dangerous and must be opposed. However, being aware of the existence and the workings of this and other psychological strategies is, paradoxically, the best we have when it comes to produce quality reflexive processes that result in truthful evaluations, to pursue morally worthy actions, as well as to embrace practical identities that are genuinely committed to morality.

---

[40] This is assumed by Korsgaard as being part of her first-person normative perspective, according to which moral obligations are generated from what *the agent herself* takes to be normative. For the critique about immoral practical identities, see Cohen, 1996, pp. 183 - 187. For Korsgaard's response in this regard, see Korsgaard, 1996, pp. 254 - 258.

[41] The precise process of how this is carried out is described by Korsgaard under the name *reflective endorsement*. Drawing on Kant, she maintains that the test of reflective endorsement is precisely the method "used by actual moral agents to establish the normativity of all their particular motives and inclinations (...) [It] is not merely a way of justifying morality. *It is morality itself*". For her argument in this regard, see Korsgaard, 1996, chapter 03: *The Authority of Reflection*.

In sum, akin to virtue signaling, the self-deception involved in self-signaling is a reinforcing attitude, closely related to the perception and moral assessment an agent makes about oneself. One's view about oneself, about one's own moral identity, and ultimately about one's very interest in morality, make SGN so powerful. Part of their success depends on whether psychological mechanisms aimed at preserving the agent's moral integrity such as self-deception are operational.

This comes to the fore when we approach the problem of adherence to green narratives from a moral standpoint. Lacking moral literacy that ensures quality in forming and assessing the moral judgments underpinning one's decisions and actions, one's practical identities (despite being shaped as a means of providing us guidance in the moral terrain) tend to paradoxically fulfill the very opposite role. They may function as shields of self-deception, hindering the agent's access to one's own maxims, character and will.

Kant's moral psychology allows us to ascertain the underlying reasons as to why this is worrisome, but most importantly, it provides us with elements to delve beyond the virtue signaling broadly construed. These elements not only disclose the normative aspect of decisions surrounding green issues, but also reveal the ethical nuances and the psychological aspects that lie at the heart of our behavior towards narratives we adhere to in spite of acknowledging them as deceptive.

Abstract. Why do agents, who consider themselves ethical or green consumers, endorse green narratives they acknowledge to be deceptive? In this paper I draw on Kant's moral psychology to propose a conceptual framework that provides us with elements to explain individual-level adherence to symbolic green narratives (SGN). From a theoretical perspective, I show how Kant's moral psychology provides insights consistent with the state of the art in contemporary social psychology regarding why green consumers fail to enact their best judgment when it comes to endorsing such narratives. My approach is complementary to other theories in arguing that many of our actions in the moral field, and particularly actions and decisions concerning green narratives, are typically pursued aiming not merely at virtue signaling but also at *self*-affirmation concerning one's own virtue.

***Keywords*:** symbolic green narratives; moral psychology; practical identity; virtue-signaling; self-deception.

REFERENCES

Adams, C. J. (2007). Caring About Suffering. *The feminist care tradition in animal ethics*, pp. 198 -226.

Adams, C. J. (2018). *The sexual politics of meat*. Routledge, pp. 548 - 557.

Alvaro, C. (2019). Is Animal Suffering Really All That Matters? The Move from Suffering to Vegetarianism. *Journal of Agricultural and Environmental Ethics*, 32(4), 633-645

Annunziata, A., Ianuario, S., & Pascale, P. (2011). Consumers' attitudes toward labelling of ethical products: The case of organic and Fair Trade products. *Journal of Food Products Marketing*, *17*(5), pp. 518 - 535.

Attah-Boakye, R., Adams, K., Yu, H., & Koukpaki, A. S. F. (2022). Eco-environmental footprint and value chains of technology multinational enterprises operating in emerging economies. *Strategic Change*, *31*(1), pp. 99 - 116.

Berthon, P. R., Ferguson, S. T. L., Pitt, L. F., & Wang, E. (2021). The virtuous brand: The perils and promises of brand virtue signaling. *Business Horizons*.

Bowen, F. (2014). *After greenwashing: Symbolic corporate environmentalism and society*. Cambridge University Press.

Blackford, R. (2021). Grandstanding: The Use and Abuse of Moral Talk. *The Philosophers' Magazine*, (92), pp. 116 - 117

Carver, C. S. & Scheier, M. F. (1998). *On the Self-Regulation of Behavior*. Cambridge University Press, New York.

Chowdhury, R. M. (2017). Emotional intelligence and consumer ethics: The mediating role of personal moral philosophies. *Journal of Business Ethics*, 142(3), pp. 527 - 548.

Cohen, G. A. (1996). Reason, humanity, and the moral law. In: C. M. Korsgaard. *The sources of normativity*. Cambridge University Press, pp. 167-188.

Crowe, R., & Simon, W. (2000). *Who are the ethical consumers?* (pp. 1- 44). Manchester: Cooperative Bank.

De Pelsmacker, P., Driesen, L., & Rayp, G. (2005). Do consumers care about ethics? Willingness to pay for fair-trade coffee. *Journal of Consumer Affairs*, 39(2), pp. 363 – 385.

Dhar, R., & Wertenbroch, K. (2012). Self-signaling and the costs and benefits of temptation in consumer choice. *Journal of Marketing Research*, 49(1), pp. 15 - 25.

Dhir, A., Sadiq, M., Talwar, S., Sakashita, M., & Kaur, P. (2021). Why do retail consumers buy green apparel? A knowledge-attitude-behaviour-context perspective. *Journal of Retailing and Consumer Services*, *59*, 102398.

Follows, S.B. and Jobber, D. (2000). Environmentally responsible purchase behaviour: a test of a consumer model. *Eur. J. Marketing* 34, pp. 723 – 746.

Gaard, G. (2002). Vegetarian ecofeminism: A review essay. *Frontiers: A Journal of Women Studies*, *23*(3), pp. 117 - 146.

Giner-Sorolla, R. (2001). Guilty pleasures and grim necessities: affective attitudes in dilemmas of self-control. *Journal of personality and social psychology*, 80(2), 206.

Goleman, D. (1985). *Vital Lies, Simple Truths: The Psychology of Self-Deception*. Touchstone, New York.

Hamilton, V. L. & Sanders, J. (1981). The effect of roles and deeds on responsibility judgments: The normative structure of wrongdoing. *Social Psychology Quarterly*, 44(3), 237.

Hainmueller, J., Hiscox, M. J., & Sequeira, S. (2015). Consumer demand for fair trade: Evidence from a multistore field experiment. *Review of Economics and Statistics*, *97*(2), pp. 242 - 256.

Herman, B. (2022). *Kantian Commitments: Essays on Moral Theory and Practice*. Oxford University Press.

Kaspar, K., Newen, A., Dratsch, T., de Bruin, L., Al-Issa, A. & Bente, G. (2016). Whom to blame and whom to praise: Two cross-cultural studies on the appraisal of positive and negative side effects of company activities. *International Journal of Cross Cultural Management*, 16(3), pp. 341 – 365.

Kitcher, P. (2003). What is a Maxim? *Philosophical Topics*, 31(1/2), pp. 215 - 243.

Kivetz, R., & Zheng, Y. (2006). Determinants of justification and self-control. *Journal of Experimental Psychology*: General, 135(4), 572.

Kant, I. (2019) [REL] *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.

_____ (2017) [MS] *The Metaphysics of Morals*, ed. by Denis, L., & Gregor, M. J. Cambridge, United Kingdom: Cambridge University Press.Press, UK.

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

_____ (2011) [REL] *Die Religion innerhalb der Grenzen der blossen Vernunft*, ed. by Höffe, O. Walter de Gruyter.

_____ (1999) [KrV] *Critique of Pure Reason*, trans. by Paul Guyer and Allen W. Wood, Cambridge: Cambridge University Press.

_____ (1997) [VE] *Lectures on Ethics*, ed. by Heath, P. & Scheenwind, J.B., Cambridge University.

Korsgaard, C. M. (2009). *Self-constitution: Agency, identity, and integrity*. OUP Oxford.

_____ (2011). Self-Constitution and Irony. In: *A Case for Irony*, ed. Jonathan Lear. Cambridge, Harvard University Press, pp. 75 - 83.

Korsgaard, C. M., Cohen, G. A., Geuss, R., Nagel, T., & Williams, B. (1996). *The sources of normativity*. Cambridge University Press.

Martinez, L. F., & Jaeger, D. S. (2016). Ethical decision making in counterfeit purchase situations: the influence of moral awareness and moral emotions on moral judgment and purchase intentions. Journal of Consumer Marketing.

Miller, T., & Maxwell, R. (2017). *Greenwashing culture*. Routledge.

Mitchell, L., & Ramey, W. (2011). Look how green I am! An individual-level explanation for greenwashing. *Journal of Applied Business and Economics*, *12*(6), pp. 40 - 45.

Moeller, H. G. (2009). The moral fool: A case for amorality. Columbia University Press.

Nyilasy, G., Gangadharbatla, H., & Paladino, A. (2012). Greenwashing: A consumer perspective. *Economics & Sociology*, *5*(2), 116.

Obermiller, C., Burke, C., & Atwood, A. (2008). Sustainable business as marketing strategy. *Innovative Marketing*, *4*(3), pp. 20 - 27.

Papish, L. (2018). *Kant on Evil, Self-Deception, and Moral Reform*. Oxford University Press.

Powers, W. T., & Powers, W. T. (1973). Behavior: The control of perception.

Shepherd, R., Magnusson, M. & Sjoden, P.O. 2005. Determinants of consumer behavior related to organic foods. Ambio. 34, pp. 352 – 359.

Smith, D. L. (2004). Why we lie: The evolutionary roots of deception and the unconscious mind. Macmillan.

Stets, J. E. & Burke, P. J. (2000). Identity Theory and Social Identity Theory, *Social Psychology Quarterly* 63(3), pp. 224 – 237.

Sticker, M. (2021). *Rationalizing (Vernünfteln)*. Cambridge University Press.

Tosi, J., & Warmke, B. (2016). Moral grandstanding. *Philosophy & Public Affairs*, *44*(3), pp. 197 - 217.

an der Werff, E., Steg, L., & Keizer, K. (2013). It is a moral issue: The relationship between environmental self-identity, obligation-based intrinsic motivation and pro-environmental behavior. *Global environmental change*, *23*(5), pp. 1258 - 1265.

Verplanken, B., & Holland, R. W. (2002). Motivated decision making: effects of activation and self-centrality of values on choices and behavior. *Journal of personality and social psychology*, 82(3), 434.

Willemsen, P., Newen, A. & Kaspar, K. (2018). A new look at the attribution of moral responsibility: The underestimated relevance of social roles. *Philosophical Psychology*, *31*(4), pp. 595 - 608.

Zanchet, M. E. 2022. Towards a Holistic View of Self-Deception in Kant's Moral Psychology. *Con-Textos Kantianos*, n.16, pp. 194 – 2019.

Zhang, L., Li, D., Cao, C., & Huang, S. (2018). The influence of greenwashing perception on green purchasing intentions: The mediating role of green word-of-mouth and moderating role of green concern. *Journal of Cleaner Production*, *187*, pp. 740 - 750.

4.3. A3 - Regarding Animals: Kant's Account of Self-Deception and Its Relevance to Animal Welfare Advocacy

Zanchet, M. E. (2022). Regarding Animals: Kant's Account of Self-Deception and Its Relevance to Animal Welfare Advocacy. *Aufklärung,* v.9, n.3, pp. 11 - 30

Aufklärung Revista de Filosofia

AUFKLÄRUNG, João Pessoa, v.9, n.3, Set.-Dez., 2022, p.11-30
DOI: https://doi.org/10.18012/arf.v9i3.65348
Recebido: 30/11/2022 | Aceito: 26/12/2022

Regarding Animals: Kant's Account of Self-Deception and Its Relevance to Animal Welfare Advocacy

[Sobre os Animais: A Visão da Kant sobre o Auto-Engano e sua Relevância para a Defesa do Bem-Estar Animal]

Maria Eugênia Zanchet *
Universität Bayreuth, Germany

Abstract: In recent decades, the deconstruction of the anthropocentric paradigm has placed issues such as animal exploitation at the heart of modern ethical and meta-ethical debates. This topic has also been the focus of attention within Kantianism. In light of the fundamental differences between humans and animals and the impossibility of assigning direct duties to animals posed by Kant's theory, making his practical philosophy useful for the defense of animal welfare seems an impossible task, other things being equal. Added to this difficulty is the criticism that a theory based on universal principles and individual duties may be inadequate to deal with animal welfare under cultural and situational considerations. By addressing these challenges, I aim to show that Kant's practical philosophy provides valuable resources for animal welfare advocacy. In this spirit, I contend that, though limited, Kant's account of self-deception is a promising way to bridge animal welfare alongside the ethos of Kant's moral theory.
Keywords: Kant's ethics; animal welfare; self-deception; moral psychology.

Resumo: Nas últimas décadas, a desconstrução do paradigma antropocêntrico colocou questões como a exploração animal no centro dos debates éticos e meta-éticos modernos. Este tema tem sido o foco de atenção também no âmbito dos estudos Kantianos. À luz das diferenças fundamentais entre seres humanos e animais e da impossibilidade posta pela teoria de Kant de atribuir deveres diretos aos animais, tornar sua filosofia prática útil para a defesa do bem-estar animal parece uma tarefa impossível. A esta dificuldade soma-se a crítica de que uma teoria baseada em princípios universais e deveres individuais pode ser inadequada para lidar com o bem-estar animal sob considerações culturais e situacionais. Ao abordar estes desafios, pretendo mostrar que a filosofia prática de Kant fornece recursos valiosos para a defesa do bem-estar animal. Neste espírito, eu sustento que, embora limitado, o tratamento de Kant acerca do auto-engano é uma forma promissora de fazer a ponte entre o bem-estar animal e o ethos da sua teoria moral.
Palavras-chave: Ética Kantiana; Auto-Engano; Bem-Estar Animal; Psicologia Moral

Introduction

In the past few decades, the deconstruction of the anthropocentric paradigm has been placing topics such as animal exploitation at the center of contemporary ethical and meta-ethical debates. Recently, human treatment of animals has started to

* *Ph.D candidate at the Department of Philosophy and Economics at the Universität Bayreuth, Germany, working on the phenomenon of self-deception in Kant's Ethics. Her research interests include moral psychology, discourse, and decision-making processes. Email: me@mariaeugeniazanchet.com*

receive due attention, including within Kantian studies[1].

Criticisms from within as well as from outside animal welfare movements capture two aspects relevant to the moral debate on the topic. The first concerns the claim of *universality* of certain ethical theories that address animal welfare. Socially progressive groups as well as critics of animal welfare movements[2] are often skeptical that drawing on ethical theories whose principles claim to be universal can do any good. That is because once exploiting animals is deemed morally wrong, assessing individuals' actions as immoral may result in hostility towards those perceived as performing immoral acts. In addition, voices within animal welfare movements raise objections to overly *individualistic* ethical theories, which, in their view, place decision-making and agency models within the constraints of an individual's standpoint, thus failing to address structural social formations and to fully contemplate the social changes needed to form more egalitarian societies.

If correct, both views lend support to the view that universal normative approaches that focus on individualistic models of agency flatten cultural nuances, as they leave little or no room for social and situational particularities that mark cultural diversity. In this spirit, progressists tend to conclude that the best strategy to address animal exploitation is to avoid adopting individualistic approaches and, instead of targeting individual deeds, one should focus on global and structural changes. For instance, one should hold accountable large corporations commodifying animals, center one's attention on public policies that guarantee animal and social welfare, and so on.

As a matter of fact, overly individualistic ethical approaches can lead to a number of issues, such as those associated with moral motivation. For instance, it has been argued that the perception that individual actions are only narrowly effective may lessen individuals' motivation to undertake actions whose ends go beyond that discrete action's scope. Chignell (2020) draws on Kant's *Religion* to argue that even when one's individual efforts to curb animal exploitation may appear ineffective, reasonable hope bolsters one's psychological resolve to take actions out of duty.

Due to the narrow scope of this paper, I do not develop straightforward answers to the problems of universality or individuality in general. Like Chignell, I bring these concerns into the framework of Kant's ethics, problematizing *to what extent* an ethics that focuses on individual duties and is grounded on universal principles can address the problem of how we treat animals. Particularly, I discuss the degree to which Kant's ethics can accommodate premises that protect animals from human harm, and argue that within Kant's theoretical framework, his account of self-deception amounts to an insightful path to take for making a compelling case towards animal welfare.

In order to do so, I proceed in two stages. In the first section, I show that, because of the way Kant structures his taxonomy of duties and most importantly, because such categorization is based on a reading about fundamental differences between humans and animals, assigning *direct* duties to animals is not a possibility[3]. However, as I argue in the second section, this is not the same as denying that certain elements of Kant's practical philosophy can provide a valuable resource for advocating for animal welfare. In that section, I provide a brief reconstruction of the self-regarding duties to strive for moral perfection and self-knowledge to argue that the duty against self-deception has an important complementary function to the former ones. To illustrate the centrality of truthfulness in ensuring their fulfillment, I resort to empirical evidence as well as to epistemic considerations that will ultimately allow us to point out the limitations of Kant's account of self-deception and of Kant's moral theory when it comes to addressing animal welfare.

I then proceed to consider the merits of Kant's ethics and conclude by arguing

AUFKLÄRUNG, João Pessoa, v.9, n.3, Set.-Dez., 2022, p.11-30

that avoiding self-deception is a fundamental step one should take in order to avoid performing actions of animal exploitation - especially when it comes to the main forms of animal abuse discussed throughout the empirical literature. I point to the idea that the criticisms regarding individualistic and universal ethics may be easily given an argumentative coating so as to be used by commonsense individuals to rationalize[4] their behavior. Finally, I address these criticisms by showing that, for Kant, individual and collective progress are operative and tied together within his account of moral progress.

By the end of this discussion, I will have built a balanced case for the helpfulness of Kant's ethics for animal welfare. On the one hand, Kant's account of self-deception seems a promising way to ground the defense of animal welfare from an objective point of view, since it establishes a criterion of objectivity that at the same time takes into account the theoretical framework underlying the attribution of merely indirect duties to animals. On the other hand, the duty against self-deception as such runs into inescapable epistemic limitations, which make the assignment of objectively valid duties to the institution of animal welfare especially difficult in Kant.

SECTION I

In the Introduction to the *Metaphysics of Morals* Kant takes the first step toward systematizing his taxonomy of ethical duties by presenting the two ends one ought to adopt as duties: one's own perfection and the happiness of others (MS 6:385). These ends underlie the division of ethical duties between *duties to oneself* and *duties to others* (MS 6:386; GMS 4:421). In turn, each of these duties bifurcates into perfect and imperfect duties. Perfect duties are also called negative or limiting duties because they prohibit or comment on specific actions. Imperfect duties, on the other hand, are called positive and are wide in scope, for they command the adoption of specific ends. (MS 6:419).

Throughout his ethical writings, Kant insists on the priority of perfect over imperfect duties. Imperfect duties allow latitude for choice and are thus subordinated to perfect ones, which do not.

Meanwhile, duties to oneself are core to Kant's ethics since they may be regarded as fundamental to our use of practical reason. By looking at Kant's stance on the consequences of violating these duties, one might find it helpful to appreciate their relevance. For instance, while violating duties to others impacts an agent's worth in specific circumstances, to violate duties to oneself *completely robs* one's inner worth (VE 27:344; MS 6:404, 420).

Although one can hardly argue for the priority of duties to oneself over duties to others[5], noting the centrality of these duties is crucial for at least two reasons. First, because it allows us to understand how Kant's view of animal cognition motivates his understanding of the moral status of animals and of the duties we have toward them, which will be discussed next. Second, because the centrality of duties to oneself directly entails the centrality of the duty against self-deception for cases involving animal welfare. This will be discussed in section 02.

Importantly, duties to oneself relate to the way one regards one's own worth, an idea that carries a key aspect of Kant's view on the moral status of animals. According to him, the principle underlying self-regarding duties rests on something *unique* to human beings, namely, freedom[6]. Although animals have the capacity for the power of choice, they lack freedom of choice, as their power of choice is determined by their sensible nature (VE, 27:344).

Moreover, although both share a sensible nature that makes them "needy beings" capable of using their power of choice to satisfy their sensible urges, only human beings are capable of elevating themselves above their own animality (KpV: 61). They are able to choose to prioritize rationality over their sensuous needs and inclinations. For instance, "someone who has a surgical operation performed on himself feels it without doubt as something bad; but through reason he and everyone declares it to be good." (KpV: 61) In other words, as human beings, we are free to manipulate our natural drives and eventually put them at the service of what is required by reason. (KU, 5:432). Action according to duty emphasizes the unique ability of human beings to value their own worth, that is, to act based on their self-regard as free rational beings.

In contrast, Kant assumes that in the absence of a rational nature, animals are driven by a comparable, but not identical principle, namely instinct. Most notably, Kant states that "[w]e can attribute to animals an analogue of reason (*analogon rationis*), which involves connection of representations according to the laws of sensibility" (VM 28: 276[7]). It poses an unbridgeable distinction between humans and animals, rooted in the way these two species' souls are constituted - rather than just a matter of degree (VM 28: 689 - 690;).

Throughout his ethics, Kant emphasizes that although part of human nature encompasses animality, human beings are animals "*endowed* with reason" (MS 6: 456, my emphasis) as the capacity for rationality is interconnected with the capacity for morality. Accordingly, the point that interests Kant in this claim is not simply that animals do not possess reason. Instead, Kant's interest seems to lie in what follows from it. Because animals lack reason (and because moral obligation requires reason), animals cannot stand under moral obligation.

Kant's account of human predispositions deepens this gap between humans and non-human animals even further. According to him, three predispositions determine human beings: animality, humanity, and personality (REL 6: 26). The predisposition to animality concerns characteristics of our sensible nature, such as self-preservation, perpetuation of the species, and preservation of one's "capacity to enjoy life" (MS 6: 420). The predisposition to humanity, also called physical self-love, involves "the inclination to gain worth in the opinion of others" (REL 6: 27), and precisely because of this, it involves a certain (sensibly limited) use of practical reason. In contrast, the predisposition to personality is described by Kant as a condition that makes human beings susceptible to *morality*. In particular, it concerns the "susceptibility to respect for the moral law *as of itself a sufficient incentive to the power of choice*" which is, in turn, rooted in "practical reason of itself, i.e., in reason legislating unconditionally" (MS 6: 28).

Critically enough, this distinction is core to the *Metaphysics of Morals*, since it is there that he develops his taxonomy of duties and addresses the consequences that follow once one violates them. For example, Kant states that some violations of duties may deprive a human of her own personality, leaving one only with animality, and as a result making one "a plaything of the mere inclinations and hence a *thing*" (MS 6: 420, my emphasis).

But while in some passages Kant claims that animals are mere "things [...] with which one can do as one likes" (ANTH 7: 127), in those cases his tone seems descriptive, suggesting a discrepancy between his use of the concept of "things"[8] to characterize animals when compared with the overly derogatory sense invoked to refer to humans who violate certain duties.

Addressing these tensions may help clarify Kant's view on the moral difference between human beings and animals. Unlike animals, whose actions are directly

determined by their sensible nature, humans do possess freedom to act in accordance with principles. According to Kant's "fact of reason" (KpV 5:29 - 30), by acting on maxims, human beings become immediately conscious of the moral principle, which forces itself irresistibly on them (REL 6: 36; MS 6: 438). Whenever an agent acts on the moral principle, she simultaneously makes good use of her practical reason. While it seems unreasonable to hold accountable beings that lack the ability to act otherwise, the same does not apply to beings who, instead of making proper use of their practical reason, give up such prerogative and choose to act driven by sensible principles. Therefore, the freedom to act according to principles and the idea that a free rational will is unique to human beings, sheds light on the rationale behind Kant's negative connotation when referring to human beings losing their personality[9]. Crucially, giving up such a prerogative is so strongly criticized by Kant because moral obligations are grounded in the human ability to act according to principles, and particularly, to moral principles. It is therefore not surprising that the moral predisposition, i.e., the predisposition to "treat [one]self and others according to the principle of freedom under laws" (ANTH 7: 322) is the highest level of distinction between human beings "as opposed to the other inhabitants of the earth"[10] and acting against such natural predisposition may compromise one's own humanity and personality, thus "degrad[ing] humanity *below the animal level"* (LE 27: 392, my emphasis).

On account of such a conception of human nature in which Kant states that human beings hold an *autonomous* will, i.e., the ability to act on a self-imposed moral law, he recommends in his *Lectures on Pedagogy* (9: 449 - 450;) a model of education that concentrates on the harmony between the natural predispositions to animality, humanity, and personality.

Notwithstanding the favorable approach to the predisposition to animality, Kant has no reservations in assigning only *indirect* duties to animals. In his view, direct moral obligation depends on a condition of relational existence, that is, it is only possible once a relationship between two or more rational beings or of a rational being with herself is in place. Part of Kant's argument is that such beings must, by virtue of their rational will, be capable of being passively or actively obligated[11] (MS 6: 443). That is, this is not to say that only rational beings have intrinsic moral value, but rather to stress the point previously raised: that moral obligation requires reason, since it depends on agents' ability to regard themselves[12] as part of a relation of obligation.

Another way to get to it is by saying that the beings within such a relationship must possess *dignity*, that is, something that by virtue of being inherent to their rational capacity, gives them absolute internal worth. (MS 6: 435). While having a price indicates that something "can be replaced with something else as its equivalent, (...) what is elevated above any price, and hence allows no equivalent, has a dignity." (GMS 4: 435). As Kant claims in the *Metaphysics of the Morals* (6: 462),

> The respect that I have for others or that another can require from me (*observantia aliis praestanda*) is therefore recognition of a dignity (*dignitas*) in other human beings, that is, of a worth that has no price, no equivalent for which the object evaluated (*aestimii*) could be exchanged.

Acknowledging the dignity in the humanity of others (and in one's own humanity) is a duty of virtue that arises from the respect one feels for the moral law in these beings. In turn, respect is a typically human moral feeling that accounts for the main subjective condition according to which rational beings are capable of experiencing morality (REL 6: 27). Further, the feeling of respect is what makes human beings capable of incorporating the moral principle in the maxim of their actions, insofar

as it is indissociable from one's awareness regarding the authority of morality (KpV 5: 80).

Animals lack the abilities for respect and, analogously, do not possess unconditional internal worth and therefore dignity in Kant's sense. This lies at the bottom of Kant's claim that only human beings can be *objects* of our respect (KpV 5: 76), as well as at the heart of the claim that animals are not able to be obligated.

In terms of the moral relationship we may have with regard to animals, two notable implications arise. First, as discussed, the duty we have *with regard to* animals is nothing more than parasitical of the self-regarding duty to promote certain moral dispositions towards one's own moral perfection (MS 6: 443; VE 27:458 - 459). Along these lines, the duties we have with regard to animals are only indirect .

AUFKLÄRUNG, João Pessoa, v.9, n.3, Set.-Dez., 2022, p.11-30

The second implication follows from the directionality of duties, in particular the idea that indirect duties are not lesser. Instead, they merely have a different directional structure than that of direct duties. Thus, due to the fact that duties with regard to animals are indirect, Kant famously forbids all kinds of animal abuse that might degrade the agent who performs it. For example, one ought not exploit animals' brute strength beyond their capacity, promote unnecessary experiments, or kill animals unnecessarily (MS 6: 443), and so forth.

However, by virtue of settling the attribution of inner worth in reason, the possibility of moving towards moral protection to animals seems highly unlikely. Assigning direct duties to animals would require that these beings do have absolute inner worth. However, within the Kantian framework, granting inner worth to animals would involve taking on counterintuitive outcomes, such as assuming that they are endowed with rational abilities and self-assessment just as human beings are, and that this in turn would make them *fitfor* or be *subjects of* moral obligation.

But pointing to this plethora of outcomes is hardly akin to claiming that Kant's practical philosophy is completely fruitless in contributing to the discussion about how we should treat animals. In what follows, I argue that Kant's account of self-deception plays a crucial role in prohibiting the core ways in which animals are abused in our times. I assess whether this account can successfully ground ethical treatment of animals. Moreover, I argue that if one is to comply with the ethos of Kant's theory while at the same time avoiding the most widespread behaviors that involve or contribute to animal exploitation, counteracting self-deception is key. Finally, I address the criticism that Kant's ethics is narrowly focused upon individuals and show how the account of self-deception applied to the treatment of animals can be insightful when it comes to collectively furthering moral progress.

Section II

In discussions of Kant's claim of indirect duties with regard to animals, it is the promotion of right dispositions toward moral perfection that is typically invoked in the literature. As discussed, such duties are indirect because they are primarily directed at human beings (ends in themselves, capable of moral obligation) but may end up indirectly affecting how we treat animals. The duty to oneself that commands one to strive for moral perfection provides what we might call a "bridge" that grants animals a certain degree of moral worth. As Kant explains, this duty has two aspects:

> [it] is narrow and perfect in terms of its quality; but it is wide and perfect in terms of its degree, because of the *frailty (fragilitas)* of human nature. It is a human being's duty to *strive* for its perfection but not to *reach* it (in this life), and his

> compliance with this duty can, accordingly, consist only in continual progress. (MS 6:446)

This duality results from the discrepancy between the ends we must strictly pursue and the limitations imposed by human nature in terms of fulfilling this moral commandment. Therefore, this duty is deemed by Kant to be perfect only in terms of its object, that is, "*with regard to* (...) the idea that one should make it one's end to realize" (MS 6:446), while subjectively, it is an imperfect duty.

Striving for moral perfection, as well as all duties to oneself, involves the duty of self-knowledge, for the latter provides a guide through the reasons for the maxims we adopt, and therefore through the moral disposition we develop. Further, within the context of duties to oneself, and despite the fact that our self-knowledge is sensibly limited, Kant recommends that one ought to fathom oneself" in terms of [one's] moral perfection in relation to [...] duty" (MS 6: 441).

But Kant is by no means naïve when dealing with human psychology. He acknowledges that the duty of self-knowledge, and thus the duty of moral perfection, is potentially hampered by tendencies to partiality in the assessment we have of ourselves, as well as wishful thinking that we have a good heart (MS 6: 441, REL 6: 42f).

In laying out the duty to truthfulness (MS 6: 429 - 431), Kant addresses these tendencies. Because it is perfect or negative, the duty to truthfulness is described in the *Metaphysics of Morals* as the duty *not to lie,* for lying is equivalent to "the *greatest* violation of a human being's duty to himself regarded merely as a moral being" (MS 6: 429, my emphasis). It is associated with another form of lying, the inner lie or (henceforth) self-deception. The violations that both forms of lying represent, Kant holds, are blameworthy in that they amount to violating the "humanity in [one's] own person."

Kant is firmly opposed to lying and self-deception insofar as they violate perfect duties to oneself. In effect, imagining some common situations of self-deception may be insightful to better appreciate the centrality that the prohibition of self-deception takes in Kant's ethics. Consider, for example, an *opportunistic carnivore,* who maintains that her eating meat is morally permissible because, given the scale of the industry that produces meat and its inability to be sensitive to individual demands, one's particular actions are not likely to increase the harm that is already done to animals in the production chain[13]. Suppose that the opportunistic carnivore additionally accepts that animals are capable of feeling pain and, furthermore, have an interest in their own existence and that, therefore, causing unnecessary harm to animals is morally wrong.[14] She however diverts her attention away from that in order to justify her practice of eating meat and thus to defuse moral problems potentially connected to her action. The failure to correctly assess her motivation (that is, the failure to be impartial when assessing the maxim behind her behavior) leads her to the idea that her action does not conflict with what morality demands. As a consequence, she can regard her action as good (or at least, not wrong) and still hold the thoughts that caused her to put the morality of that action into question in the first place. By obfuscating the possibility that her deed might be against duty, moral perfection is, for that agent, out of the question.

However, such a case is not merely anecdotal. Many of the actions that (in industrial societies) are performed against animals stem from rationalizations[15] whose justifications are reduced to what is described in the literature as "4Ns",[16] that is, in terms of the view that eating meat is *natural*, *normal*, *necessary,* and *nice*. According to Joy (2010), the ideology of animal abuse is built upon rationalizations concerning it being *natural*, on account of what is supposedly written into our biology as humans; *normal*, because it is a widely accepted social practice; and *necessary*, by virtue of

animal meat being portrayed as a requirement for human health. The fourth N justification (*niceness*) is further described by Piazza et al (2015) in a set of studies as referring to the hedonic pleasure related to the consumption of animal meat.

Additionally, some individuals embrace justifications more obviously connected to their alleged ethical concern. Some are discussed under the name of local consumption (Pilgrim, 2013), conscientious omnivorism (Rothgerber, 2015), as well as justifications based on the defense of the virtues involved in (ethical) meat consumption (Scruton, 2004).

Yet, with this body of empirical evidence in hand, one is in a better position to understand how the duty against self-deception relates to animal welfare. To do so, consider the following epistemic profiles: a) *Naïve individuals* are agents with little or no awareness of a specific moral problem who, for situational or constitutive reasons (lack of abilities) are largely incapable of such awareness; b) *Commonsense agents*, those with little or relative awareness of a specific moral problem under normal intellectual conditions of decision making belong to this model[17]; and c) *Committed agents* are, in turn, all agents with relatively high awareness of the moral problem, which amounts to a privileged epistemic position. An additional way to further lay out this distinction is to say that while naïve agents' reasoning is hindered by external causes, commonsense and committed agents are, in contrast, not situationally hindered by complex reasoning, being able to, for example, delve deeper into moral problems, compare different ethical traditions in order to determine the best course of action, and so forth.

Consider now that agents of all three profiles perform actions that cause harm to animals, such as, for example, meat eating.Is it reasonable to maintain that in all three cases, the agents are self-deceived, i.e., that they rely on the rationalizations described above to justify their moral transgressions regarding animals? The answer to this question is *no*.

Provided with information that conflicts with their beliefs on the immorality implied once one causes harm to animals, commonsense agents might resort to any of the rationalizations just discussed as a way to make meat eating acceptable in their own eyes. They could, for instance, appeal to the supposed need for animal protein to then claim that "You cannot get all the protein, vitamins, and minerals you need on an all plant-based diet" (Piazza et al. p. 118). It is thus plausible to claim that this would be a case of self-deception where the agent shifts her attention away from, say, factual data on nutrition in order to evade the responsibility to take a moral stance should she come to firmly draw the conclusion that eating animals is morally wrong.

The case of committed agents can also be reasonably described as involving self-deception, once, for example, an agent might engage in exceptionalism in order to act against her better judgment. She might claim that refraining from harming animals is indeed a universal moral requirement but that for situational reasons (e.g., the agent is diagnosed with anemia, which supposedly requires the consumption of animal protein), such a requirement does not apply to her own case[18].

It seems, however, that such attitudes are able to be maintained, meaning, they are allowed to co-exist alongside one's striving for moral perfection. In contrast, fulfilling the duty against self-deception would prevent such rationalizations as it would guide the agent towards a more unbiased assessment as well as thoughtful scrutiny concerning the reasons behind one's actions with regard to animals. As a perfect duty, it would be even more authoritative than the duty to increase one's own moral perfection, as it outlaws all forms of untruthfulness the agent might engage in when assessing her actions toward animals. Unlike moral perfection, which admits different degrees and

may even be obtained in different aspects of the moral agent's life (e.g., someone may seek moral perfection concerning this or that aspect of her life), the duty prohibiting self-deception would be operative for most cases of animal abuse in which the agents' rationality is not externally hindered (as in cases of naïve individuals). Within the context of the treatment of animals, the duty against self-deception would be relevant for a significant number of cases (since most of the justifications that people typically use to justify their meat eating fit into the 4Ns).

However, whereas the distinction between the three profiles highlights where self-deception succeeds,[19] it also points to an important limitation of Kant's theory. It concerns whether the success of the account of self-deception in Kant's defense of animals depends on the agent's degree of self-awareness. As a matter of fact, that is precisely what prevents one from attributing self-deception to explain naïve agents' immoral behavior. One may object that Kant's recommendation in favor of an enlightened life circumvents this limitation. In the same line that he states "not to avoid the places where the poor who lack the most basic necessities are to be found but rather to seek them out", in other passages, Kant insists that "man has a duty to raise himself from the crude state of his nature, from his animality (...) toward humanity (...) [and] to *diminish his ignorance* by instruction and to correct his errors." (MS 6:387, my emphasis). However, placing the argument's strength into such reasoning is rather unconvincing. One would still need to take at least some degree of *prior awareness*, for instance, knowing that sick people exist in the first place, or getting some idea about their own ignorance, and then on how to diminish it, should one have the moral resolve to expose oneself to such situations.

This limitation is worth further elaboration, for it gets even more nuanced when accompanied by the criticisms stated at the beginning of this paper, i.e., that a moral theory overly focused on individuals would flatten social differences, stigmatizing both peoples and individuals whose culture incorporates animal exploitation. At the heart of such criticism is the implicit assumption that the traditions of each culture have absolute intrinsic worth, and therefore Kant's allegedly individualistic theory would bear poor resonance against those practices.

Regardless of how we deal with this apparent limitation, it must be emphasized that this criticism itself tends to rely on self-deception[20]. Notably, while naïve individuals would hardly be in an epistemic position to endorse such a reasoning, commonsense individuals often take an interest in it, by partaking in the naïve profile, feeling exempted from moral obligations towards animals. This would represent instances of self-deception whereby (1) commonsense individuals might voluntarily "remain naïve" regarding topics they anticipate to prompt moral obligations[21], as well as cases where (2) these individuals instrumentalize naivety in order to feel justified in moral transgressions.

In the first case (1), the fact that someone resists the attempt to revise one's beliefs provides evidence that such an agent no longer counts as naïve, for the very attempt reveals one's clear interest in morality. Certainly, one might be *truly* naïve, but given the increasing availability and growing presence within the media of topics related to the ethics of animal abuse, the question then concerns how long one can remain truly naïve. As soon as she is aware that her previously taken-for-granted behavior may be morally problematic, an agent may qualify as what I am calling here a commonsense individual. As she strives to remain naïve, say, by persistently turning a blind eye to a problem's ethical standing or to considerations that would otherwise render such a problem worthy of moral regard, that agent deceives herself not only about her own beliefs, but primarily about her relationship and interest in morality.

In the latter case (2), it is not uncommon to see meat eaters[22] leaning on traditional folk practices as a way to justify their diet habits or their lack of attention to animal welfare. For them, animal welfare movements are to be accused of proselytism resulting from "first-world" privileges. Behind their argument is not only the misleading idea that plant-based diets are less affordable than traditional ones, but above all, that education, i.e., access to knowledge (and with it, leaving the scope of naivety) is a "first-world" trait. These individuals take the premise that for some people (indigenous communities, people from certain regions or social classes) preventing animal exploitation is simply not an option and extend it to themselves.

But the tendency to create justifications does not only apply to commonsense individuals striving to mimic naivety. Further behavior typical of individuals who want to evade moral demands and moral awareness may surprisingly lean in the opposite direction. This is exemplified by localist movements or individuals who identify themselves as vegetarians but eat meat. For example, localist arguments revolve around the idea that being aware of the origin of one's meat provides sufficient justification to render permissible one's eating meat.

Analogous to the case of commonsense individuals seeking to appear naïve, committed agents create contingent criteria they know they can meet to justify their transgressions. Free range, humane slaughter or occasionalism (when someone consumes meat only on certain occasions but still identifies with the anti-animal exploitation cause) illustrate some of these criteria. The difference is rather that such criteria involve attitudes in which agents often recognize the moral transgression as such, as opposed to the cases of commonsense agents trying to emulate naivety discussed earlier. But despite the differences, in both cases self-deception seems to lie in the same rationale, that is, in enabling those agents to violate their own moral conscience. By establishing and then meeting criteria contingent on that moral problem, these agents create false excuses that justify their moral transgressions.

From what has been described, it can be seen that self-deception is an extremely valuable idea to explain the persistence and pervasiveness of moral transgressions that are to some extent acknowledged as such[23]. Self-deception is present in situations where, faced with a moral problem, the agent offers justifications that mask their real motivation and that evince their interest in morality, while simultaneously rendering permissible in their own eyes what they perceive as potential moral transgressions.

However, it is important to note that self-deception does not arise in the case of truly naïve subjects, nor does the duty against self-deception. For Kant, once we know that a problem is a *moral* problem, we must act accordingly. But if we have no clue that a certain problem is worthy of moral consideration, then the obligation against self-deception does not emerge. This is particularly evident for cases of fake news and deep fakes, that is, cases where fabricated lies resist an agent's critical scrutiny. Of course, as a moral agent, one has a duty to scrutinize and revise one's moral beliefs. Yet, when it comes to some forms of fake-news but especially to deep fakes it may be argued that such agents are thrown back into the group of naïve individuals, insofar as those untruths bear the power to twist reality and completely obscure the moral relevance of certain ethical issues. It may seem that this is a weak limitation, since the ultimate success of fake news often relies on echo chambers, confirmation bias mechanisms, etc[24]., against which individuals would have a duty to be alert. However, reality distortions may be more pervasive than one would first imagine. This is epitomized by ideologies themselves, which essentialize certain beliefs, holding them virtually beyond the scrutiny of agents who are part of societies in which those ideologies hold sway. Furthermore, such ideologies impact the progression of moral and scientific knowledge,

for example backing biased research and outcomes, which in turn reinforce the prevailing beliefs. Consider for example the fallacious ideas that brain differences between men and women justify behavioral discrepancies, or that eating meat and dairy is crucial for being healthy. These naturalized ideas, advanced both scientifically and in the field of popular knowledge, disregard for instance how evidence on brain functioning was historically gathered (Rippon, 2019) as well as the existence of alternative and healthier nutrients' sources[25].

To sum up, the existence of naïve individuals imposes an explanatory hindrance by which self-deception, and hence the duty against self-deception, fails to arise. This in turn stands as a limitation of Kant's account of self-deception with respect to animal welfare, as well as a range limitation of his moral theory to the animal question, because along with its non-applicability to a certain group of individuals, it fails to provide objective and universal elements for animal advocacy.

It is important to draw attention to the fact that such a limitation of scope is ultimately a weak one. This is because in terms of determining the moral community to whom a moral theory is addressed, one can reasonably maintain that no moral theory fares better in this respect. This is to be said that such a limitation is weak because it applies to any moral theory, since no moral theory is universally applicable to all scenarios. In other terms, moral theories typically assume that there are agents for whom moral problems do not emerge, such as children, comatose individuals, or cognitively impaired persons. While weak, this limitation certainly counts as a constraint as it narrows the scope of validity of Kant's framework of self-deception for dealing with cases of individual decision-making about animal matters.

A possible way to address such a limitation would be raising public awareness to moral problems, so as to progressively decrease the number of naïve individuals. However, while the limitation of Kant's self-deception account applies most clearly to individuals belonging to this epistemic profile, it is not restricted to that. This is because immoral deeds might result from self-deception but also from one's explicit rejection of striving for one's own moral perfection. Consider for example cases of *akrasia*, i.e., cases of agents who acknowledge the immorality of their actions and yet perform them thoughtfully. In such cases, self-deception *might* arise *post hoc*, to diminish the discomfort brought about by the moral transgression by means of shifting its moral burden (e.g., 'eating meat is immoral, it would be even worse if the animal had not been free-ranged, therefore, consuming meat from free-range animals is acceptable'). Yet, there are situations where a complete and utter disregard for the morality of particular actions might take place. In such cases, the nuisance and therefore the cognitive strategy of dealing with the moral discomfort does not arise. Although both cases imply the interdiction (to a greater or lesser degree) of self-perfection, in neither of them does self-deception *necessarily* occur. Thus, the ability of Kant's account of self-deception to deal with animal welfare problems runs up against not only the weakness of the will of particular individuals in performing moral actions, but also in the practice of clear-headed and unconcerned immoral actions.

That granted, further limitations[26] are claimed against Kant's ethics in general. Along with the one just discussed, there is a further broader claim according to which Kant's theory, supposedly over-focused on individuals' duties, actions and moral dispositions, tramples on more collective social considerations and creates situations that potentially blame individuals. While Kant's account of self-deception provides us with a plausible explanation concerning the *nature* of such criticism (meaning, that the criticism itself is based on rationalizations), its content also needs to be addressed. However, before I engage in that, I will briefly highlight at least two points according to

which Kant's practical philosophy and his account of self-deception may be fruitful when it comes to our moral treatment of animals.

Firstly, it is particularly important to mention that the way Kant grounds his theory on the primacy of reason has encouraged reactions that point to the inconsistency of using such an "intellectual gap" as a rationale for ascribing moral worth. The incongruence poses problems such as problems of scope[27]; leading to counterintuitive[28] or ruthlessly exploitative[29] positions. The emergence of those problems suggest the need for a significantly stronger philosophical reassessment of concepts (such as *animal dignity*), key to our understanding of animal moral standing. Such an undertaking was carried out by many Kantian philosophers[30], who advanced their disagreements with Kant on animal welfare and, as a result, advanced theories tackling the gaps left by Kant.

Yet, what is usually overlooked concerning the set of rationalizations one engages to justify one's moral transgression, is that it indicates that *we do have reasons* to engage in animal abuse. For example, one may ground one's choices in aesthetic reasons (pleasure-oriented reasons), economic reasons (related to the cost of dietary choices), as well as social reasons (associated with avoiding social norms confrontation), and so forth. In the latter case, these reasons are often associated with what Joy (2010) calls *carnism*, that is, a belief system that promotes structural animal abuse while simultaneously discouraging its members from challenging it.

This leads to a second point that I believe makes Kant's theory especially fruitful in advancing the possibility of animals partaking in morality as legitimate subjects of our moral concerns. Since most harm caused to animals is unnecessary[31], one may claim that, in general, animal exploitation is performed based on prudential reasons[32], meaning that the maxims of these actions are happiness or, in Kant's words, self-love (rather than duty) oriented. Acting from reasons typically linked to the satisfaction of one's sensible desires or, in most cases, to the lack of moral maturity to claim ownership over one's agency and think for oneself, hardly cohere with Kant' ethos.

Contra this view, one could contend that there are cases in which actions out of duty are not required, since wide duties allow "a latitude (*latitudo*) for free choice in following (complying with) the law" (MS 6: 390). In fact, within Kant's theory, it is not an issue that part of our actions are merely prudential as, incidentally, striving for happiness is a natural fact of our nature[33]. However, in the following lines Kant asserts that latitude refers to the *priority* one gives to certain maxims (for, as discussed, duties of virtue prescribe ends rather than discrete actions). Therefore, it is fundamentally problematic that an agent, upon identifying a moral problem, creates exceptions for oneself in order to allow acting from self-love rather than duty. In this respect, and paradoxically, despite the limitation it sets, Kant's account of self-deception is crucially insightful. To say that we tend to ground on prudential reasons the performance of deeds that we, to some extent, acknowledge as moral transgressions, amounts to claiming that we perform actions whose maxims lack truthfulness[34]. That is, by rationalizing our reasons for action, we incur self-deception, instrumentalize the suffering of animals but most importantly, fail to regard ourselves as worthy of dignity and respect as well as an end in ourselves.

Yet, instrumentalizing animal suffering out of prudential reasons seems importantly inconsistent with the Kantian ethos regarding virtuous agency[35]. Virtuous agents, for Kant, are those who perform actions which have morally worthy ends. Under ideal conditions, they acknowledge actions that demand a moral stance and, in striving for moral perfection, act accordingly. As autonomous agents, we must

> in *all our maxims* [...] not forget to be submissive to [reason], not to detract from it in any way, and not to curtail in any way—*through delusion based on self-love*—the authority of the law (even though our own reason gives it) by positing the determining basis of our will, even if in conformity with the law, still in something other than the law itself and in respect for this law. (KvP: 82, my emphasis)

From this notion, Kant idealizes a "kingdom" of autonomous beings, that is, "a systematic *union of various rational beings* through common laws" (GMS 4: 433, my emphasis), in which its members are legislating members who regard themselves as such, and who "carefully" (GMS 4: 463) navigate moral territory guided by duty.

The centrality of this concept[36] points to an important reply to the criticism according to which Kant's theory is overfocused on individuals, namely that, for Kant, moral progress is inscribed in the idea of community.

Implied in the idea of the kingdom of ends is the view that, within a moral community, the way its members use their freedom must be consistent among them, both internally and externally. This means that their will must constrain but also be constrained by the will of others, as well as internally by what morality requires. Put another way, these agents must be able to autonomously incorporate moral principles into their maxims over self-interested principles.

Such a task is, however, not without difficulties. This is because acting on incentives other than moral principles (such as, the satisfaction of one's own passions, interests, one's own happiness, and so on) amounts to an inescapable propensity that belongs to human nature. If this were an easy task, our society would not witness the persistence of ideologies of human and animal exploitation, let alone the rationalizations discussed above. Moreover, acting in such a way amounts to a tendency that is collectively self-encouraged, that is, one in which society's members tend to stimulate and even to reward. Take for example the connection between supremacist ideologies of gender and species, discussed by Adams (2018). Societies in which such ideologies are prevalent tend not only to favor but also to reward practices that subscribe to such ideologies. Examples are behaviors that associate and at the same time denigrate and objectify women and animals, or that connect meat consumption and violence against animals to traits such as virility, pleasure, and strength.

In *Religion* (6: 42f), Kant insists that in performing actions out of interests other than morality, one tends to engage in self-deception, *viz.*, "to lie to oneself in the interpretation of the moral law to its prejudice". As a natural tendency reinforced among agents, and moreover, while opposing autonomy, self-deception represents a hindrance to morality. Therefore, forming a moral community may be in the best interest of autonomous and virtuous moral agents. As Kleingeld (1995, p. 160, my translation) puts it, a community whose agents perform actions out of duty

> aims at the moral improvement of all mankind, at the preservation of morality and the counteraction of evil. (...) As long as they are not united in such a way, people will seduce each other to evil (...) [due to their] innate inclination to evil. (...) The goal of an ethical community is to remove this obstacle to morality. Because the highest good, as a community of the virtuous and as a moral world, requires such a community, it is a duty to join an ethical community.

Thus, it is clear that Kant does not reject culture and collectivity. His practical philosophy rather emphasizes that moral progress amounts to an end that must be *collectively* pursued[37]. Crucially, Kant's practical philosophy, through its concepts of autonomous agency, the duties towards moral perfection and self-knowledge, but especially because of its prohibition of self-deception, is in position to accommodate the

view that in societies oriented towards moral progress, their members are entitled to make cultural demands.

But although the self-regarded duty against self-deception ultimately supports cultural demands for a decent treatment of animals within Kant's ethics, Kant's practical philosophy remains to be reclaimed to ensure that it is not misconceived and thus used to support rationalizations and further actions that prioritize selfish interests (be those individual or institutional) or self-interested principles.

AUFKLÄRUNG, João Pessoa, v.9, n.3, Set.-Dez., 2022, p.11-30

Conclusion

At the core of the Kantian ethos is the idea that all human beings should pursue the formation of an ethical community, that is, a kingdom of ends to be achieved by means of its members' moral behavior. That said, and with respect to how we should treat animals, two things are granted.

The first concerns the twofold function of reason. While the rational faculty precludes animals from direct duties, it also points to duties that result in the ethical treatment of animals. The self-regarding duties to strive for moral perfection and self-knowledge already drive the human will towards an ethical treatment of animals. However, as such, these are not completely efficient, as they do not prevent one from accessing one's actions in benefit of one's own self-interest, e.g., disfavoring the moral law. In this regard, the perfect duty against self-deception backs up the agent's duty to pursue moral perfection, since it ensures truthfulness regarding how the agent accesses one's maxims. As a result, when applied to how we treat animals and alongside the duty to strive for moral perfection, Kant's approach to self-deception prevents actions grounded in the agent's self-interest as well as in one's failure to properly review one's own beliefs.

Second, the very postulation of a kingdom of ends further stresses the social nature of Kant's theory, for even individual moral progress is a component of collective moral progress. The approach Kant takes in laying out the framework of his moral theory necessarily binds individual deeds (maxims of discrete actions) and the character of particular agents to a conception of moral progress that is, by definition, social.

Objections to Kant's theory as a viable approach to animal ethics because of its emphasis on individual duties are therefore misguided. Additionally, the universal character of Kant's ethics, which relies on the distinctions he draws concerning human and animal nature, serves ultimately a narrow purpose. For example, in the *Casuistry of the Metaphysics of Morals* Kant addresses a series of discrete cases, taking into account situational considerations that eventually emerge. Moreover, the very existence of imperfect duties, i.e., duties that command broadly, demonstrates that there is ample ground for discretionary maxims and actions.

In this setting, the importance of the universality of duty against self-deception becomes even more pronounced. This perfect duty provides a feasible compass for how one assesses one's maxims, one's striving for moral perfection and self-knowledge, as well as one's own character. Critically, absent the universal validity of this duty and the (universal) considerations of human nature, the path to moral relativism gathers its momentum.

However, even though Kant's account of self-deception appears to be a promising approach to improving animal welfare, it ultimately falls short in providing a convincing justification for the ethical treatment of animals. This is because, taken individually, the duty against self-deception stumbles upon significant epistemic

constraints, as previously discussed. The current scenario of epistemic unreliability stemming from the increase in reality-falsifying technologies underscores the validity of such a critique.

Therefore, someone who is interested in this perspective might ask: how does a theory whose ethos rests on collective moral progress come to be practically relevant to animal welfare? In this paper, I discussed the role of the duty against self-deception in addressing the question at hand. Notwithstanding the epistemic limitations it bears, in the context of Kant's ethics, and especially when paired with the duties towards moral perfection and self-knowledge, the duty against self-deception may be regarded as an effective touchstone for commonsense and committed agents who may need practical guidelines on how to navigate the global issue of animal exploitation.

References

Allen, N. E., Key; T. J., Appleby P. N.; Travis R. C.; Roddam A. W.; Tjønneland, A.; Riboli, E. Animal foods, protein, calcium and prostate cancer risk: the European Prospective Investigation into Cancer and Nutrition. *British Journal of Cancer*, 98(9), (2008): 1574–1581.

Adams, C. J. The sexual politics of meat. In *Living with Contradictions*. Routledge. (2018): 548–557

Almeida, M. and Berstein, M. 'Opportunistic Carnivorism', *Journal of Applied Philosophy*, 17 (2), (2000): 205–11.

Bernecker, S., Flowerree, A. K., & Grundmann, T. (Eds.). *The epistemology of fake news*. Oxford University Press, 2021.

Callanan, J. & Allais, L. (Eds.). Kant and Animals. Oxford: Oxford University Press, 2020.

Camenzind, S. Der Paratext in Immanuel Kants Metaphysik der Sitten und seine (tier-)ethischen Implikationen. In: Tierethik transdisziplinär: Literatur—Kultur—Didaktik. Bielefeld: Transcript, (2018): 43–59.

Chignell, A. Hope and Despair at the Kantian Chicken Factory: Moral Arguments about Making a Difference, in: Lucy Allais & John J. Callanan (eds.), *Kant and Animals*. Oxford University Press, (2020): 213–238.

Cunningham, A. Lymphomas and animal-protein consumption. *The Lancet*, 308(7996), (1976): 1184–1186.

Bauman, Z. Does ethics have a chance in a world of consumers? Harvard University Press, 2009.

Bilewicz, M.; Imhoff, R.; Drogosz, M. The humanity of what we eat: Conceptions of human uniqueness among vegetarians and omnivores. *European Journal of Social Psychology*, *41*(2), (2011): 201–209.

Descartes, R. *The Philosophical Writings of Descartes: Volume III: The Correspondence, In:* Cottingham, J., Stoothoff, R., Murdoch, D., & Kenny, A (Transl). Cambridge University Press, 1991.

Dowsett, E.; Semmler, C.; Bray, H.; Ankeny, R. A.; Chur-Hansen, A. Neutralising the meat paradox: Cognitive dissonance, gender, and eating animals. *Appetite*, *123*,(2018): 280–288.

Engel Jr, M. The commonsense case for ethical vegetarianism. *Between the Species*, *19*(1), 1, 2016.

Golob, S. What do animals see? Intentionality, objects, and Kantian nonconceptualism. *Kant and animals*, in: Lucy Allais & John J. Callanan (eds.), *Kant and Animals*. Oxford University Press, (2020): 66-88.

Joy, M. *Why we love dogs, eat pigs, and wear cows. An introduction to carnism*. San Francisco, CA: Red Wheel/Weiser, 2010.

Hartung, T. Opinion versus evidence for the need to move away from animal testing. *Altex*, *34*(2), (2017): 193-200.

Herman, B. We are not alone: A place for animals in Kant's ethics. In Kant on persons and agency. Cambridge: *Cambridge University Press*.(2018): 174–191.

Herzog, H. *Some we love, some we hate, some we eat. Why it's so hard to think straight about animals*. New York: Harper Perennial, 2010.

Kant, I. *Lectures on Ethics*, Heath, P.; Scheenwind, J.B.; (Eds.), Cambridge, United Kingdom:

Cambridge University Press, 1997.
Kant, I. *Critique of the Power of Judgment*, P. Guyer (Ed.), Cambridge, United Kingdom: Cambridge University Press, 2000.
Kant, I. *Lectures on metaphysics*. Cambridge, United Kingdom: Cambridge University Press, 2001.
Kant, I. *Critique of Practical Reason*, W. S. Pluhar (Ed.), Cambridge, United Kingdom: Cambridge University Press, 2002.
Kant, I. Lectures on pedagogy (1803) In: I. Kant (Author); R. Louden; G. Zöller (Eds.), *Anthropology, History, and Education*. Cambridge, United Kingdom: Cambridge University Press,(2007): 434-485.
Kant, I. *Lectures on Anthropology*. Cambridge, United Kingdom: Cambridge University Press, 2012.
Kant, I. *Groundwork of the Metaphysics of Morals: A German–English Edition*. Gregor, M.; Timmermann, J. (Eds.), Cambridge, United Kingdom: Cambridge University Press, 2014.
Kant, I. *The Metaphysics of Morals*, Denis, L.; Gregor, M. J. (Eds.). Cambridge, United Kingdom: Cambridge University Press, 2017
Kant, I. *Religion within the boundaries of mere reason and other writings*, Wood, A. W.; Giovanni, G. D.; Adams, R. M., (Eds.) Cambridge, United Kingdom: Cambridge University Press, 2019.
Kleingeld, P. *Fortschritt und Vernunft: Zur Geschichtsphilosophie Kants*. Königshausen und Neumann, 1995.
Korsgaard, C. M. Fellow creatures: Kantian ethics and our duties to animals. The tanner lectures on human values, 2004.
Korsgaard, C. M.. Kantian ethics, animals, and the law. Oxford Journal of Legal Studies, 33(4), (2013): 629-648.
Korsgaard, C. M.. A Kantian Case for Animal Rights. In: Margot Michael, Daniela Kühne & Julia Hänni (eds.) Animal Law - Tier and Recht: Developments and Perspectives in the 21st Century 3-27. Zurich: Dike Verlag, 2012.
Leach, S.; Piazza, J.; Loughnan, S.; Sutton, R. M.; Kapantai, I.; Dhont, K.; Douglas, K. M. . Unpalatable truths: Commitment to eating meat is associated with strategic ignorance of food-animal minds. Appetite, 171, 105935, 2022
Loughnan, S.; Haslam, N.; Bastian, B. . The role of meat consumption in the denial of moral status and mind to meat animals. *Appetite*, 55(1), (2010): 156–159.
McLear, C. Kant on animal consciousness. *Philosophers Imprint*, 11(15). University of Nebraska - Lincoln, (2011): 01–16.
Miller. H. *Jury convicted two activists today of Break-and-Enter and mischief, acquitted a third, for their role in exposing animal cruelty*. Canada News Media. Available at: , 2022, July 11. Last access on: July 19, 2022.
Müller, N. D. Kantianism for Animals: A Radical Kantian Animal Ethic, 2022.
Molfetta, M.; Morais, E. G.; Barreira, L.; Bruno, G. L.; Porcelli, F.; Dugat-Bony, E.; Minervini, F. . Protein Sources Alternative to Meat: State of the Art and Involvement of Fermentation. *Foods*, 11(14), 2022.
Mourouti, N.; Kontogianni, M. D.; Papavagelis, C.; Plytzanopoulou, P.; Vassilakou, T.; Psaltopoulou, T.; Panagiotakos, D. B. Meat consumption and breast cancer: a case–control study in women. *Meat Science*, 100, (2015): 195–201.
Pilgrim, K. 'Happy cows,''happy beef': a critique of the rationales for ethical meat. *Environmental Humanities,3* (1), (2013): 111–127.
Pinheiro Walla, A. Kant's moral theory and demandingness. *Ethical Theory and Moral Practice*, *18*(4), (2015): 731–743.
Pinheiro Walla, A. Happiness in Kant's Practical Philosophy: Morality, Indirect Duties, and Welfare Rights, 2022.
Potter, Nelson. Kant on duties to animals. *Jahrbuch für Recht und Ethik* 13, (2005): 299–311.
Rothgerber, H. Can you have your meat and eat it too? Conscientious omnivores, vegetarians, and adherence to diet. *Appetite*, *84*, pp. 196-203.Scruton, R. (2004). The conscientious carnivore. In S. Sapontzis (Ed.), *Food for thought*. Amherst, NY: Prometheus, (2015). 81–91.
Sticker, M. *Rationalizing (Vernünfteln)*. Cambridge University Press, 2021.
Timmermann, J. When the tail wags the dog: animal welfare and indirect duty in Kantian ethics. Kantian review, 10, (2005): 128–149.
Williams, N. M. Affected ignorance and animal suffering: Why our failure to debate factory farming puts us at moral risk. *Journal of Agricultural and Environmental Ethics*, *21* (4),

(2008): 371–384.
Wolk, A. Potential health hazards of eating red meat. *Journal of Internal Medicine*, 281(2), (2017): 106–122.
Wood, A. W. (1998). Kant on duties regarding nonrational nature: Allen W. Wood. In *Aristotelian Society Supplementary Volume*, University College London: The Aristotelian Society, Vol. 72, No. 1, (1998): 189–210
Wood, A. W. *Kant's ethical thought*. Cambridge: Cambridge University Press, 1999.
Wyrębska-Đermanović, E. Kantian moral education for the future of humanity: The climate change challenge. *Journal of Philosophy of Education*, *55*(6), (2021): 1045-1056.
Zanchet, M. E. *The Moral Psychology of Individual-level Adherence to Symbolic Green Narratives*. Estudos Kantianos. Forthcoming 2023.

AUFKLÄRUNG, João Pessoa, v.9, n.3, Set.-Dez., 2022, p.11-30

NOTAS

1 See Korsgaard (2004, 2013, 2014). Wood (1998); Timmermann (2005); Herman (2018); Potter (2005); Callanan & Allais (2020); Camenzind (2018); Müller (2022).

2 Marxist-inspired progressive groups that do not identify the utility of the ethical dimension to social debate tend to criticize animal advocacy movements as superficial and elitist. Many rely on the widespread idea that material problems lack an equally historical and material analysis, thereby leaving little or no room for moral theories. While it is agreed that Marx does not reject ethics as a whole, it is arguable that he opposed theoretical models of which Kantianism is a part. For an informative discussion of the similarities between Kant's political theory and Marx's historical materialism, see also Wood (1999, pp. 226 - 244). Others find support in the distortion of postmodern social theories, such as Zygmunt Bauman's (2009). In this scenario, inadequate interpretations of Bauman's critique of the overfocus on individual responsibilities becomes ammunition for arguments against ethics focused on individual duties.

3 All things equal. For once one challenges central assumptions such as that the animals' worth is purely derivative and that, therefore, it is permissible to treat animals as mere means to our ends, such assumptions become hardly plausible. In this regard, see Camenzind (2021, pp. 07 - 08). Similarly, "to say that duties to abstain from cruel treatment of animals are not duties *to* them" (Potter, 2005, p. 33) seems particularly problematic.

4 See also KU 5: 464f.

5 In fact, for his time, Kant seems to hold a progressive view on animals in comparison to other philosophers. For example, unlike the mechanistic conception typically attributed to Descartes (Cottingham, 1978), according to which animals are described as mere automata, Kant does not deny that animals are capable of feeling pain or making decisions. For textual evidence on Descartes' views concerning animals, see especially Descartes' correspondences (Cottingham, J. et al., 1991) to Plempius for Fromondus (1637: 414 - 415) and to Reneri for Pollot (1638: 06). In correspondence to Mersenne (ATIII:85), Descartes denies that animals are capable of feeling pain.

6 The view I advance here is in line with Kant's position that natural inclinations in themselves are not culpable (REL 6: 58), since the very idea of moral transgression is only possible through rationality (CON 8: 115; REL 6: 37). Moreover, it coheres with the idea that animality is a fact of human nature which no one can ever overcome (VE 27:441).

7 In the *Metaphysics of Morals*, Kant terms "conscience" the moral (and intellectual) predisposition and states that it is inescapable to all rational beings. Unsurprisingly, conscience underlies what differentiates humans and animals, given that the latter are capable of associative abilities, perceiving objects, or intuiting particulars (Golob, S. 2020), but do not have inner sense (VM 28:277) or a sufficiently sophisticated consciousness that allows them to reflect on their behavior or on themselves (ANTH 7:153). For an in-depth account of this aspect, see McLear (2011).

8 In this respect there is a clear parallel between moral and legal obligations, both of which involve, for Kant, the possibility of reciprocal constraints, since those who obligate and those who are obliged must be capable of entering into this relationship in the first place. As summed up by Pinheiro Walla (2022, p. 141), such a "formal structure ensures that no one is bound by another in a way she could not, at least in principle, bind her in return".

9 I outline this example based on Almeida and Bernstein's (2000) consequentialist description of the *opportunistic carnivore*. People who, from a consequentialist point of view, see no moral impediment to meat consumption "suggest that a switch to carnivorism would not,

in itself, result in any additional detriment to the welfare of any animal" (p. 205). In my interpretation, these agents are self-deceived to the extent that they favor a twisted consequentialist interpretation of reality in order to accommodate justifications for a matter they actually see as morally problematic (since, for example, they do not deny that vegetarianism is required in nonindustrial production scenarios).

10 Cf. Engel Jr. (2016).

11 The pervasiveness of direct and indirect rationalization strategies is drawn by *Piazza et al* (2015) from the findings concerning the insidious persistence of the so-called "meat-paradox," a paradox faced by agents who display a tension between their thought and actions of harming animals.

12 Although the studies refer exclusively to meat consumption, I consider this classification sound for the present purposes for three general reasons. First, because meat eating is *the* representative of human exploitation towards animals. Not only does it represent an enormous fraction of human abuse toward animals, but also, *qua* a food habit, it relates to a pervasive attitude in people's daily lives. Second, because the so-called "meat-paradox" is the best documented animal exploitation phenomenon in the literature (Bilewicz, 2011; Dowsett et al, 2018; Herzog, 2010; Joy, 2010; Loughnan et al, 2014), thus providing empirical corpus as well as documented qualitative and quantitative practical evidence for the present discussion. Third, because the justifications used to make meat consumption permissible are typically the same when it comes to further actions against animals, such as clothing, entertainment, etc. Additionally, because, as Joy (2010) and Piazza et al (2015) point out, the first three justifications (naturality, normality and necessity) are involved in a variety of historical practices such as slavery and sexism, which in turn provides us with interesting insights into the magnitude of the use of psychological mechanisms with moral implications, such as self-deception.

13 In this context, self-deceptive actions are to be distinguished from admittedly *akratic* ones in that in the former, the agent seems to retain a certain interest in morality and resorts towards self-deception as a way to preserve her moral integrity, whereas in cases of *akrasia*, agents would be in a position to acknowledge their failure to act in accordance with their better judgment.

14 I will address this criticism more squarely at the end of this section.

15 Since access to certain information may trigger one's moral radar, moral agents who wish to avoid facing these moral considerations might choose to engage in a sort of *strategic* (Leach, S. et al., 2022) or *affected* ignorance. For a discussion regarding the moral dimension of affected ignorance as well as the practical applications to animal welfare, see Williams (2008).

16 Here I do not refer to the commonsense epistemic profile *per se*, but to the typical Western meat eater who does have material means as well as access in terms of knowledge to switch to a diet free of animal exploitation.

17 In this regard, see Cunningham et al, (1976); Allen et al, (2008); Wolk, A. (2017); Mourouti et al (2015); Molfetta et al (2022).

18 Significantly, the chief general limitation of Kant's theory is that it insufficiently ensures what is *currently discussed* as an ethical treatment of animals, and that by virtue of this, it becomes obsolete (Camenzind, 2021, p. 06). In particular, that it fails to grant animals integral ethical considerations (rather than merely parasitic on human interests). In this context, it is relevant to mention that although "the legislatures in Austria (1988), Germany (1990), and Switzerland (2003) have decreed that animals are, nominally, not things anymore (corresponding articles were laid down in Austria in the Allgemeinen Bürgerlichen Gesetzbuch (ABGB § 285a) in 1988, in Germany in the Bürgerlichen Gesetzbuch (§ 90a BGB) in 1990, and in Switzerland in the Zivilgesetzbuch (Art. 641a Abs. 1 ZGB) in 2003)" (Camenzind, 2021, p. 06), the accompanying legal regulations are still under contention. The fact that the largest part of atrocities against animals (livestock industry) still takes place under the law is evidence of that. For a striking example of how the prevailing legal approach still favors an economic system based on animal abuse, see the case of animal rights activists convicted after live-streaming unlawful animal handling (Miller, 2022). Yet, despite the disagreements over such an understanding, it is striking that concerning the practical implications to animal welfare, Kant's theory not only differs but also comes up short when contrasted with rival traditions. This point is raised by Camenzind (2021, p. 05) who points out that, for instance, Reagan's utilitarian theory "leads ineluctably to veganism, the total abolition of invasive animal experimentation, and comprehensive dissolution of animal agriculture and recreational hunting, because these

practices violate animals' right to bodily integrity, or their fundamental right to life".

19 A problem of scope arises from the difficulty of establishing differences between animals and persons whose rational capacities are hampered strongly enough to justify attributing indirect duties in regard to animals while direct duties to humans. According to Potter (2005, p. 305), taking Kant's view on animals prima facie would ultimately prevent the attribution of direct duties to "severely retarded [humans], those in a permanent vegetative state, those who are permanently comatose, irreversibly demented or senile, born anencephalic, and so forth."

20 Potter, 2005.

21 Wood, 1998.

22 In this regard, I especially highlight the works of Korsgaard (2004, 2012, 2013) and Möller (2022).

23 It does not seem possible, however, to extend this claim to assert that the entirety of animal exploitation is practiced for prudential reasons via self-deception to the extent that actions of agents of the naïve epistemic profile would not belong to this group. This sharply narrows down the success of Kant's account of self-deception in protecting animals.

24 Although Kant assumed that the pursuit of happiness is a fact of human nature, he did not identify happiness with human beings' natural end (GMS 4:396), assigning the title of human beings' "true vocation" (GMS 4:390) solely to *the good will*, which is the product of reason and thereby holds unconditional worth.

25 Note that Kant's recommendation in MS (6: 443) is that instrumentalizing and abusing animals is blameworthy, yet by virtue of the agent's duties to oneself. What I am proposing here is different in that it challenges Kant's own view.

26 The idea of a *kingdom* of ends is core to Kant's ethical theory, as the variation of the categorical imperative's third formula illustrates: "Act in accordance with the maxims of a universally legislative member of a merely possible *realm of ends*" (GMS 4:439, my emphasis).

27 It is difficult to argue that one's duties to oneself should be given priority over one's duties to others without leaving aside the exegetical commitment to the Kantian system of moral duties in general. However, it is worth noting that in some passages, Kant claims that there are cases where duties to oneself can take priority over duties to others, insofar as one's personality is a condition for claiming that one's actions comply with the law (VE 27:341; MS 6:417).

28 The idea that rational nature and rational will amount to the ultimate source of unconditional value is often called *Kantian logocentrism* (Wood, 1998). According to this view, human beings are to be distinguished from animals in that the former are capable of freely choosing maxims that are objectively determined by reason.

29 Furthermore, when it comes to morality, self-consciousness plays an important role. One must be aware of one's own worth, *viz*., one must regard oneself as an end if one wants to conform with self-regarding duties. Put another way, human beings *exist as* and *are able to* regard themselves as ends in themselves. Conversely, animals possess no inner worth; they exist "only as means, and not for their own sakes, *in that they have no self-consciousness*" (VE 27: 459, my emphasis).

30 It is worth noting that what I here refer to as common sense refers strictly to an epistemic profile, and does not parallel the category of common human reason discussed by Kant in the first section of GMS.

31 The stronger claim that "all harm caused to animals is unnecessary" is disputable. Some argue that medical and scientific progress vindicates animal harm. On the other hand, evidence that findings from *in vivo* studies can lead to misleading conclusions due to the physiological differences between humans and animals points to an end to the use of animals for scientific purposes or, at the minimum, to a reduction stemming from the development of alternative, safer and more effective methods (Hartung, 2017).

32 By rationalizing is meant an inappropriate use of the rational faculties, typically meaning in the moral sphere, privately or publicly presenting justifications with an eye towards the permissibility of certain actions. It refers to a flawed use of reason. For a systematized account of the term, see Sticker, 2021, pp. 08 – 10.

33 It is worthy to note that moral actions are especially pronounced in the present time. This implies that agents are often confronted with recurring moral situations, thus having the opportunity to reassess their moral stance. In such scenarios, both prior and *post-hoc* justifications merge to reduce the cognitive dissonance caused by the mismatch between one's moral perception and the performance of transgression. Rather than offering a reasonable description of how self-deception operates along the lines of social psychology,

Kant' s framework provides a methodology that allows the agent to assess one's attitudes toward animals based on a moral conscience, as well as a coherent system of duties that ultimately aims at moral perfection. In *The Moral Psychology of Individual-level Adherence to Symbolic Green Narratives* (2023, forthcoming), I discuss how Kant's framework may be clarifying when set alongside descriptive approaches.

34 It goes without saying that the fundamental problem with immoral maxims is that they are not universalizable. However, what I draw attention to here is that even universalizability crucially relies on truthfulness. For cases of self-deception, this becomes especially clear, as the agent might allow herself to violate morality by universalizing her maxims out of insincere premises.

35 See Bernecker et al. (2021), especially Ch. 9, *Echo chambers, Fake news, and Social Epistemology*.

36 Significantly, one must appreciate the part moral education plays in accomplishing such a social end. Kant's conception of moral education explicitly encompasses a view of collective moral progress that not only includes, but most importantly, extends beyond individual progress. This point is rightly raised by Wyrębska-Đermanović (2021), who argues that Kant's approach to moral education may help to promote the necessary individual changes, as well as to inform collective practices that address global issues such as climate change.

37 Meaning that it plausibly explains human behavior and at the same time provides a guideline for how one ought to act in moral matters in general, but particularly when faced with decisions involving animal exploitation.

AUFKLÄRUNG, João Pessoa, v.9, n.3, Set.-Dez., 2022, p.11-30