Studia Kantiana

REVISTA
DA SOCIEDADE
KANT BRASILEIRA

Volume 21
Número 1
Abril de 2023
ISSN impresso 1518-403X
ISSN eletrônico 2317-7462

UFPR
UNIVERSIDADE FEDERAL DO PARANÁ

Web site

www.sociedadekant.org/studiakantiana

Endereço para correspondência

[Mailing address]:

Joel Thiago Klein
Departamento de Filosofia, UFPR
Rua Dr. Faivre, 405, - 6º Andar – Ed.
D. Pedro II, Curitiba- PR
CEP: 80060-140

E-mail: studia.kantiana@gmail.com

Conselho editorial

STUDIA KANTIANA

Volume 21
Número 1
Abril de 2023
ISSN impresso 1518-403X
ISSN eletrônico 2317-7462

Artigos
[articles]

Resenhas
[reviews]

Action irrationality, systemic practical irrationality, and the remedy in Kant

[Irracionalidade da ação, irracionalidade prática sistêmica e o remédio em Kant]

Jean-Christophe Merle[1]

Universität Vechta (Vechta, Alemanha)

DOI: 10.5380/sk.v21i1.91472

Abstract

In contrast to the well explored criteria of practical rationality in Kant's philosophy, there is hardly any attempt to systematically examine the sources of practical irrationality in Kant. Yet, the development of the realm of reason among human beings depends on this core issue of his philosophy. Although Kant does not provide any comprehensive analysis of irrationality, one can distinguish three kinds of problems and sources of practical irrationality, related respectively to each of the three main dimensions of practical rationality: instrumental, practical (in a narrower sense), and systemic. Inquiring into these dimensions, this paper also explores the reasons why a lack of rationality remains even in individual actions out of duty, i.e. out of practical reason, and why only acting under an idea of systemic rationality makes possible the full achievement of rationality.

Keywords: irrationality; systemic irrationality; instrumental irrationality; practical irrationality; realm of reason.

1 Jean-Christophe.Merle@uni-vechta.de

1. Three main dimensions of practical rationality

In contrast to the criteria and conditions of practical rationality in Kant's philosophy that have been intensively explored in the last decades, there is hardly any attempt to systematically examine the sources of practical irrationality in Kant. Nonetheless, there is not only plenty said about practical irrationality by Kant, but it is also at the core of his philosophy, on which depends the development of the realm of reason among human beings, and the implementation of what Onora O'Neill calls "constructions of reason".

Although Kant does not provide any comprehensive analysis of irrationality, one can distinguish in these abundant materials three kinds of problems and sources of practical irrationality, related respectively to each of the three main dimensions of practical rationality: instrumental rationality, practical rationality (in a narrower sense), and systemic irrationality.

Instrumental rationality consists in adopting suitable means to one's ends. Its usual criteria – efficiency, regularity ("lawfulness" (*Gesetzmäßigkeit*) in Kant), consistence – are present in Kant's practical philosophy. *Practical rationality in a narrower sense* pertains to appropriate end setting. Kant does not determine it primarily by criteria of the rational choice theory – such as completeness and transitivity –, but instead as choosing ends that are suitable to the specific nature of human beings. Non-human living beings behave in accordance with laws and with an end, yet neither choosing nor having a representation of them/it. In fact, according to Kant, "only a rational being has the capacity to act in accordance with the representation of laws, that is, in accordance with principles, or has a will" (GMS, AA 4: 412).

Instinct does not provide non-human living beings only with an end/ends, but also with the means suitable to those ends. They do not really choose the means: instinct cares both for the means and for the ends. Non-human living beings behave in accordance with an end in two regards: individually, striving for self-conservation, as well as collectively, striving for the reproduction of their species. Human beings are not given their collective organization by nature. Instead, it results from individual and collective choices. Thus, human collective organization is the object of *systemic rationality*.

2. Three sources of practical irrationality

Now, there are three kinds of problems or sources of irrationality related respectively to each of these dimensions of rationality: problems or factors of irrationality pertaining (1) to the appropriate means, (2) to end setting, and (3) to collective organization.

(1) Concerning *instrumental rationality*, Kant asserts:

> Whoever wills the end also wills (insofar as reason has decisive influence on his actions) the indispensably necessary means to it that are within his power. This proposition is, as regards the volition, analytic, for [...] the imperative extracts the concept of actions necessary to this end merely from the concept of the volition of this end (GMS, AA 4: 417).

Since a human being may pursue more than one end at a time, one should talk of "actions fitting to this end" rather than "actions necessary to this end". Yet, from a non-normative point of view, a descriptively neutral characterization of the human will would be: Whoever wills the end also wills what he *believes* to be the indispensably necessary means to it that she or he *believes* to be within her or his power. Here lies a cognitive source of the irrationality of some actions, that is, wrong belief(s) about one the following objects: (i) about the indispensable – or the most appropriate – means, and (ii) about whether these means are within one's power. In the absence of reason's "decisive influence on his actions", the human being may be, or is, likely to

act irrationally, from an instrumental point of view, on the basis of one or both of these kinds of wrong belief(s). However, under reason's "decisive influence on his actions" too, such wrong beliefs may occur and lead to instrumental irrationality. In fact, Kant observes that "reason itself does not operate instinctively, but rather needs attempts, practice and instruction in order gradually to progress from one stage of insight to another" (IaG, AA 8: 18).

This road of gradual progress is paved by intermediary steps that are, from an instrumental point of view, irrational to various extents.

(2) Let us now look at the factors of *irrationality pertaining to end-setting*. Inside the same person, ends given by sensibility to the faculty of desire, that is, inclinations, which are "habitual sensible *desire*[s]" (Anth, AA 7: 452), collide with one another. This situation unavoidably results in either indeterminacy of the will ("peevish wish": Anth, AA 7: 251) or frustration of the will, to the extent to which the will is determined by inclinations:

> Now, it is impossible for the most insightful and at the same time most powerful but still finite being to frame for himself a determinate concept of what he really wills here. If he wills riches, how much anxiety, envy and intrigue might he not bring upon himself in this way! If he wills a great deal of cognition and insight, that might become only an eye all the more acute to show him, as all the more dreadful, ills [*Übel*] that are now concealed from him and that cannot be avoided, or to burden his desires [*Begierden*], which already give him enough to do, with still more needs. If he wills a long life, who will guarantee him that it would not be a long misery? If he at least wills health, how often has not bodily discomfort kept someone from excesses into which unlimited health would have let him fall, and so forth. (GMS, AA 4: 418)

Kant's point is not to deplore the existence of inclinations as habitudes. Inclinations originate in the faculty of desire, which is a natural disposition, so that it and they cannot be eradicated. Inclinations are not damaging as long as they are subordinated to reason. Thus, already the child's inclination to avoid pain ought to be directed by means of coercion into the habit to obey rules:

> children are sent to school initially not already with the intention that they should learn something there, but rather that they may grow accustomed to sitting still and observing punctually what they are told, so that in the future they may not put into practice actually and instantly each notion [*Einfälle*] that strikes them. [...] the human being must be accustomed early to subject himself to the precepts of reason (Päd, AA 9: 442).

Yet, there exist two kinds of end-setting by the faculty of desire that cannot be subordinated to reason, because they either have an excess or a deficit of habitude: (i) affects, that is, "surprise through sensation, by means of which the mind's composure (*animus sui compos*) is suspended" (Anth, AA 7: 252) and (ii) passions, that is, "inclination[s] that can be conquered only with difficulty or not at all by the subject's reason" (Anth, AA 7: 251).

Now, despite these permanent and partly deep-rooted collisions among inclinations whenever they are not subordinated to reason, human beings strive for the impossible satisfaction of all of them:

> all people have already, of themselves, the strongest and deepest inclination to happiness because it is just in this idea that all inclinations unite in one sum. [...] and yet one can form no determinate and sure concept of the sum of satisfaction of all inclinations under the name of happiness (GMS, AA 4: 399).

Whereas the "sum of satisfaction of all inclinations" is inconsistent, the satisfaction of all inclinations that are subordinated to reason is possible, to the extent to which they are subordinated to reason:

> In the sensibly practical too, reason goes from the general to the particular according to the principle: not to please one inclination by placing all the rest in the shade or in a dark corner, but rather to see to it that it can exist together with the totality of

> *all* inclinations (Anth, AA 7:251).

Yet, at the same time, passion hinders the ordering task of reason: "Inclination that prevents reason from comparing it with the sum of all inclinations in respect to a certain choice is *passion* [*Leidenschaft*] (*passio animi*)" (Anth, AA 7: 266).

Yet, since ideas originate in the faculty of reason in its speculative use, irrationality in end-setting is not possible without reason, not only because no rule and no habitude can be acquired without reason, but also because reason is also the decisive source of the idea of happiness. Kant insists on this point:

> Passion always presupposes a maxim on the part of the subject, to act according to an end prescribed to him by his inclination. Passion is therefore always connected with his reason [...]. (Anth, AA 7: 266)

Thus, reason is both the source of rationality and the decisive source of irrationality. This results from the double characterization of reason by Kant (i) as the combination of the power of universal rules and aims and (ii) as a power that "knows no boundaries to its projects" (IaG, AA 8: 18f.). Not sensibility, but rather both of these aspects of reason are the main source of irrationality in general, and the main source of irrational inclinations in particular:

> [...] it is the property of reason that with the assistance of the power of the imagination it can concoct desires not only *without* a natural drive [*Naturtrieb*] directed to them but even *contrary to* it, which desires in the beginning receive the name of *concupiscence* [*Lüsternheit*]; but through them are hatched bit by bit, under the term *voluptuousness* [*Üppigkeit*], a whole swarm of dispensable inclinations, which are even contrary to nature (MAM, AA 8: 111).

(3) By the factors of *irrationality pertaining to collective organization*, I mean the problems related to the collision between individual wills. Reason, not as the power of rules and universal aims, but reason as the power that knows no boundaries to its projects, in other words, the faculty of illimitation, gets activated and dominates the contacts between individuals:

> the inclination to freedom as a passion [...] is the most violent inclination of all, in a condition where he [the human being] cannot avoid making reciprocal claims [*Ansprüche*] on others (Anth, AA 7: 268).

Kant enumerates the forms adopted by this passion for freedom: "These passions are the *manias for honor*; *for dominance*, and *for possession* [*Ehrsucht*, *Herrschsucht*, *Habsucht*]" (Anth, AA 7: 270).

Thus, not only happiness as the idea in which "all inclinations unite in one sum" (GMS, AA 4: 399) cannot be reached because of a collision, that is because of the collision between the inclinations of the same individual, but also that the passion for freedom necessarily leads to collisions – namely, to collisions between individual wills – that likewise necessarily impede the attainment of all of its aims.

Because of these problems, (conditional) imperatives, i.e., the imperative of skill and the imperative of prudence, seem paradoxically at the same time a source of irrationality and as originating in rationality.

All *imperatives of skill* – that is all hypothetical, problematic or technical imperatives – originate in a certain way from rationality, or, more precisely, from *instrumental rationality*, which consists in adopting laws that connect suitable means to one's ends. In fact,

> All imperatives are expressed by an *ought* and indicate by this the relation of an objective law of reason to a will that by its subjective constitution is not necessarily determined by it (a necessitation) (GMS, AA 4: 413).

The development of all-purpose skills is an end of reasons, and the object of a duty toward oneself. Thus, "as a rational being he necessarily wills that all the capacities in him be

developed, since they serve him and are given to him for all sorts of possible purposes" (GMS, AA 4: 424).

Now, passions entail the use of instrumental rationality:

> Since [these] passions [*Ehrsucht*, *Herrschsucht*, *Habsucht*] are inclinations that aim merely at the possession of the means for satisfying all inclinations which are concerned directly with the end, they have, in this respect, the appearance [*Anstrich*] of reason; that is, they aspire to the idea of a faculty connected with freedom, by which alone ends in general can be attained. Possessing the means to whatever aims one chooses certainly extends much further than the inclination directed to one single inclination and its satisfaction (Anth, AA 7: 270).

Now, the factors of irrationality pertaining to collective organization may originate not only from the cognitive source of the instrumental irrationality mentioned above, which is a problem internal to instrumental rationality, but also from the separability of the instrumental rationality from the rationality concerning end setting. Indeed, Kant affirms: "Whether the end is rational and good is not at all the question here [...]" (GMS, AA 4: 415), which contrasts with Aristotle's view on prudence, because Aristotelian prudence concerns the way in which goods – not any kind of ends – are strived for, i.e., the way virtues apply "it is not possible to be good in the true sense of the word without prudence, or to be prudent without moral goodness" (NE 1144b31).

Not only all applications of the *imperative of skill*, but also the – hypothetical, assertoric and pragmatic – *imperative of prudence* originate in a certain way from rationality, in a negative regard as well as in a positive regard. In a *negative regard*, Kant states that unhappiness may motivate to violate the moral law of reason, and thus that unhappiness should be avoided for the sake of practical reason in the narrow sense:

> To assure one's own happiness is a duty (at least indirectly); for, want of satisfaction with one's condition, under pressure from many anxieties and amid unsatisfied needs, could easily become a great *temptation to transgression of duty* (GMS, AA 4: 399).

For this reason, not caring for one's own happiness must be considered as a factor of irrationality pertaining to collective organization. In a *positive regard*, although the happiness of the person acting morally is not the objective of an imperative, it is an object of hope resulting from morality itself. The first general imperative of duty commands "[...] act as if the maxim of your action were to become by your will a universal law of nature" (4:421). Actually, what Kant means here is *the* universal law of nature. Thus, one should act as if the maxim of one's action were to become fully successful. In fact, Kant affirms:

> morality alone, and with it, the worthiness to be happy, is also far from being the complete good. In order to complete the latter, he who has not conducted himself so as to be unworthy of happiness must be able to hope to partake of it (KrV, B841).

Admittedly, a hope is not the same as an imperative, and making an imperative out of happiness understood in an unconditioned way, i.e., not connected to morality, is immoral and irrational. Yet, this abuse results (through a paralogism) from a rational premise, which is that the perspective of success ought to be entailed in the will.

As is well known, the imperative of prudence is also a source of irrationality, not only because of its inconsistence, but also because this inconsistence manifests itself in the fulfillment of this imperative. In fact, any of the inclinations that are to be satisfied within the total satisfaction of all inclinations under the name of happiness (GMS, AA 4: 399) may become motivation to deviate from the pursuit of happiness. Kant describes this weakness of the choice pursuing happiness:

> the precept of happiness [i.e. the imperative of prudence] is often so constituted that it greatly infringes upon some inclinations. [...] Hence, it is not to be wondered at

> that a single inclination, determinate both as to what it promises and as to the time within which it can be satisfied, can often outweigh a fluctuating idea, and that a man – for example, one suffering from gout – can chose to enjoy what he likes and put up with what he can since, according to his calculations, on this occasion at least he has not sacrificed the enjoyment of the present moment to the perhaps groundless expectation of a happiness that is supposed to lie in health (GMS, AA 4: 399).

3. Why does a lack of rationality remain in individual action out of duty, i.e., out of practical reason?

Might the *categorical imperative*, like the imperative of the skill and the imperative of prudence, both originate from rationality and yet also be a source of irrationality? Paradoxically, it seems to be the case, for the following reason.

The second imperative of duty asserts the end of rational action: "*So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means*" (GMS, AA 4: 429).

Here, humanity refers to what is specifically human, i.e., to what does not belong to animality: reason. Thus, the second imperative of duty requires that reason be considered in each person at the same time as an end and as a means. There are frequent misunderstandings related to this second imperative of duty: (i) considering each person as an end in itself and (ii) considering reason (or persons) only as end(s) in itself/themselves. Concerning the latter frequent misunderstanding, one should not forget the expression "not only", since this expression implies that there are further ends than reason itself. Indeed, reason may be used as a means for (a) satisfying inclinations or individual wills, whether an other's or mine, (to the extent that they do not violate the moral law), as well as, for (b) promoting reason's own complete and purposive development as the end of this natural disposition.

I begin with the case (a). Reason is not only allowed by morality to be used as a means for realizing both my will or other's wills, but it is also required to be used in this way in the context of moral duty:

> Providing for oneself to the extent necessary just to find satisfaction in living (taking care of one's body, but not to the point of effeminacy) belongs among duties to oneself. [...] beyond *benevolence* in our wishes for others (which costs us nothing) how can it be required as a duty that this should also be practical, [...]? – Benevolence is satisfaction in the happiness (well-being) of others; but beneficence is the maxim of making other's happiness one's end, and the duty to it consists in the subject's being constrained by his reason to adopt this maxim as a universal law (TL; AA 6: 452).

Now, there is a moral problem resulting from benevolence: the recipient's dependence on the benefactor, gratefulness being both a duty and a dependence toward the benefactor. To the latter extent, benevolence seems to collide with the duty of respect toward others and toward oneself. Indeed, this dependence may result in the recipient's distrust and even ingratitude toward the benefactor, and possibly in the benefactor's self-distrust. Kant solves the problem raised by this possible collision through (i) requiring beneficence to be exerted anonymously (neither should the benefactor know who is the recipient nor should the recipient know who is the benefactor, i.e., beneficence should be organized collectively), and (ii) requiring (moral) friendship to be limited to communicating one's ideas, cognitive, moral and esthetic judgments, expressly excluding beneficence among friends (Merle, 2021). Both requirements pertain to what Kant calls respect toward others. Indeed, generally speaking, according to Kant, the individual moral subject ought to observe an equilibrium between beneficence (also called love) and respect:

> The chief division [of duties to others] can be that into duties to others by performing which you also put others under obligations and duties to others the observance of which does not result in obligation on the part of others. Love and respect are the feelings that accompany the carrying out of these duties. [...] they are basically always united by the [moral] law into one duty, only in such a way that now one duty and now the other is the subject's principle, with the other joint to it as accessory (TL, AA 6: 448).

This results in each person's moral behavior towards others having to observe an equilibrium between attraction and repulsion by others. Now, repulsion or respect toward others amount to restricting the use of reason as a means for (a) promoting reason's own complete and purposive development as the end of this natural disposition as well as for (b) satisfying inclinations or individual wills, whether other's or mine, that do not violate the moral law. Thus, this equilibrium raises the following two problems pertaining to social irrationality.

(a) This equilibrium lead to a limited contribution to the promotion of reason, hence to a deficit of rationality:

> It is a duty to oneself as well as to others not to *isolate* oneself [...] but to use one's moral perfections in social intercourse (*officium commercii*, *sociabilitas*). [...] forming part of an all-inclusive circle of those who, in their disposition [*Gesinnung*], are citizens of the world – not exactly in order to promote as the end what is best for the world but only to cultivate what leads indirectly to this end: to cultivate a disposition of reciprocity – agreeableness, tolerance, mutual love and respect (TL, AA 6: 473).

The restrictions related to the contribution to what is "best for the world" are expressed in this quotation by "not exactly" and "indirectly", and they are clearly related to "mutual love and respect" between which an equilibrium has to be observed. That this "duty not to isolate oneself," consisting in "cultivating mutual love and respect," and an equilibrium between both amounts to no more than a limited contribution to the promotion of reason is confirmed by the fact that moral law and the categorical imperative point to a more comprehensive end of social rationality.

Let us examine this more comprehensive end of social rationality. In the second imperative of duty, "humanity, whether in your own person or in the person of any other" ought to be an end. Now, reason is a natural predisposition that cannot be fully developed in the individual, so that reason as an end can lie only in the species:

> *All natural predispositions of a creature are determined sometime to develop themselves completely and purposively* [...] *In the human being* (as the only rational creature on earth), *those predispositions whose goal is the use of his reason were to develop completely* [Nisbet: "could be fully developed"; "*sollten sich [...] völlständig entwickeln*"] *only in the species, but not in the individual* (IaG, AA 8: 18).

As is well-known, according to Kant, the complete and purposive development of human predispositions cannot be achieved but through the succession of an "immense" number of generations:

> nature perhaps needs an immense series of generations, each of which transmits its enlightenment to the next, in order finally to propel its germs in our species to that stage of development which is completely suited to its aim (IaG, AA 8: 19).

The explanation provided by Kant in IaG for such high number of generations is neither only, nor mainly, due to cognitive obstacles against the scientific inquiry into nature and against the development of technics, but to social obstacles. The duty to observe an equilibrium between beneficence and respect is based on the statement of the existence of such obstacles. If these obstacles were not taken into account, the moral law would not result in this duty of equilibrium between beneficence and respect that is at the core of Kant's *Doctrine of Virtue*.

Yet, strictly speaking, in this matter, not the *categorical imperative* itself, is the source of the huge deficit of rationality mentioned above, but this deficit appears first in duties derived from

the categorical imperative, i.e., the "duty to oneself as well as to others not to isolate oneself" or the duty to observe an equilibrium between benevolence and respect in the relationship to others. This points to the fact (i) that, even when a person acts out of duty, i.e., motivated by reason alone, which includes the duty to develop one's talents, this individual action alone cannot lead to the full development of reason and rationality, and (ii) that, even if each person individually acted morally, these actions alone could not lead to the full development of reason and rationality. The reason for this impossibility lies in a second problem pertaining to social irrationality, which is the following one.

(b) The – real or latent – collisions among the wills of the individual moral subjects remains unresolved by the equilibrium between benevolence and respect that limits the individual wills. Now, this constant tension among the wills is the source of passions that are socially irrational. Indeed, each individual will, led by instrumental reason, seeks to extend its skills resulting in comparison to others regarding their respective skills, in competition and in fear. In turn, a motive or temptation to act immorally against others (without beneficence or without respect) is generated. This is *a permanent temptation for social irrationality*:

> It is not the instigation [*Anreize*] of nature that arouses what should be called the *passions* [*Leidenschaften*], which wreak such great devastation in his [the human being's] originally good predisposition. His needs are but limited, and his state of mind is providing for them moderate and tranquil. He is poor (or considers himself so) only to the extent that he is anxious that other human beings will consider him poor and will despise him for it. Envy, addiction to power, avarice [*Habsucht*, *Herrschsucht*], and the malignant inclinations associated with these, assail his nature, which on his own is undemanding, as soon as he is among human beings (RGV, 6: 93f.).

Thus, drawing on an analogy with Kant's theses on perpetual peace, the mere individual observance of the categorical imperative amounts more to a cease-fire than to a true peace among individuals. Generally speaking, full rationality cannot be achieved by individual rational moral actions alone. Thus, if any, the achievement of full rationality can only be the achievement of systemic rationality.

4. The full achievement of rationality through acting under an idea of systemic rationality

In Kant, this systemic rationality rests upon the thesis of the general will inspired by Rousseau. Thus, let us first have a short look at what Rousseau's concept of a general will consist in? Rousseau presents it as follows:

> There is often a considerable difference between the will of all and the general will: the latter looks only to the common interest, the former looks to private interest, and is nothing but a sum of particular wills; but if, from these same wills, one takes away the pluses and the minuses which cancel each other out, what is left as the sum of the differences is the general will (Jean-Jacques Rousseau, *Social Contract*, Book II, Ch. 3).

The core of Kant's idea of the kingdom of ends is the same as the core of Rousseau's concept of a general will despite the fact that the latter relates to external legislation, that is, to the relationships between individual free wills in their external dimension ruled by legal law, whereas the former relates to internal legislation, that is, to their relationships in their internal dimension ruled by morality or virtue. Kant explains:

> *kingdom of ends*. By a *kingdom* I understand a systematic union of various rational beings through common laws. Now since laws determine ends in terms of their universal validity, if we abstract from the personal differences of rational beings as

> well as from all the content of their private ends we shall be able to think of a whole of all ends in systematic connection (a whole both of rational beings as ends in themselves and of the ends of his own that each may set himself), that is, a kingdom of ends, which is possible in accordance with the above principles (GMS, AA 4: 433).

This quotation does not only explain what the idea of a kingdom of ends means. It also explains why this idea is consistent, hence realizable, unlike, for example, the inconsistent idea of happiness as the sum of the satisfaction of all our inclinations.

Now, the idea of a kingdom of ends provides the basis for the third imperative of duty. The imperatives of duty, the addressee of which is the individual person, are ordered according to the categories of quantity (see KrV, B106): (1) unity (or the law of nature), (2) plurality (of the persons), (3) totality (of the kingdom of ends). Totality is the unity of plurality. Accordingly, the third imperative of duty expresses in the most adequate way the systemic rationality of the categorical imperative, while entailing both two first imperatives of duties, even though the presence of the wording of the second imperative of duty appears more clearly. Kant explains:

> For, all rational beings stand under the *law* that each of them is to treat himself and all others *never merely as means* but *always at the same time as ends in themselves*. But from this there arises a systematic union of rational beings through common objective laws, that is, a kingdom, which can be called a kingdom of ends (admittedly only an ideal) because what these laws have as their purpose is just the relation of these beings to one another as ends and means. A rational being belongs as a *member* to the kingdom of ends when he gives universal laws in it but is also subject to these laws. He belongs to it as a *sovereign* when, as lawgiving, he is not subject to the will of any other (GMS, AA 4: 433).

As is well known, here again Kant draws on Rousseau despite the fact that Rousseau deals with external legislation, whereas Kant relates to internal legislation. Indeed, Rousseau explains:

> [...] each individual, by contracting, [...] finds himself engaged in a two-fold relation: namely, as member of the Sovereign toward private individuals, and as a member of the State toward the Sovereign (Jean-Jacques Rousseau, *Social Contract*, Book I, Ch. 7).

With "that each of them is to treat himself and all others never merely as means but always at the same time as ends in themselves", Kant almost expressly repeats the second imperative of duty:

> *So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means* (GMS, AA 4: 429).

What is new in the third imperative of duty, as compared to the second imperative of duty, is the point of view adopted by the individual moral agent: the point of view of the systematic union of rational beings. In other words, the person cannot be moral nor fully rational except by thinking herself or himself as acting as part of – or under the idea of – the systematic union of rational beings. Only the idea of the united or general will makes human beings capable of adequate end-setting, that is, rational end-setting. Not only social irrationality, but also individual irrationality, cannot be remedied but with systemic practical rationality. This leads Kant to sketch the steps of the education of individual human beings as follows:

> In his education, the human being must therefore 1) be *disciplined*. To be disciplined means to seek to prevent animality from doing damage to humanity, both in the individual and in society. [...]
>
> 2) The human being must be *cultivated*. [...] It is the procurement of skillfulness. The latter is the possession of a faculty which is sufficient for the carrying out of whatever purpose. [...]
>
> 3) It must be seen that the human being becomes *prudent* also, well suited for human society [...]. This requires a certain form of culture, which is called *civilizing*. [...]

> 4) One must also pay attention to *moralization*. The human being should not merely be skilled for all sorts of ends, but should also acquire the disposition [*Gesinnung*] to choose nothing but good ends. Good ends are those which are necessarily approved by everyone and which can be the simultaneous ends of anyone. (Päd, AA 9: 449f.)

The general will of the kingdom of ends cannot adopt ends apart from such "good ends [that] are those which are necessarily approved by everyone and which can be the simultaneous ends of anyone." The idea of the kingdom of ends, which is the last step of Kant's educational curriculum, contains instrumental rationality, practical rationality (in a narrower sense) and systemic rationality. Thus, in Kant's view, the idea of the kingdom of end constitutes the ultimate and only full remedy for problems or factors of irrationality pertaining to (1) the appropriate means, (2) the end setting, and (3) collective organization.

Bibliographic References

KANT, I. *Cambridge Edition of the Works of Immanuel Kant*. Paul Guyer and Allen W. Wood (ed.). Cambridge: Cambridge University Press, 1996–.

KANT, I. *Gesammelte Schriften*. Ed. Königlich Preussische Akademie der Wissenschaften (and successors). Berlin: de Gruyter (and predecessors), 1900–.

MERLE, Jean-Christophe. 'Von den Tugendpflichten gegen Andere: Liebe, Achtung und Freundschaft'. In: Jean-Christophe Merle and Carola Freiin von Villiez (Eds.), *Zwischen Rechten und Pflichten – Kants ‚Metaphysik der Sitten'*, Berlin: De Gruyter, pp.223-245.

ROUSSEAU, Jean-Jacques: 'On the Social Contract'. In: *The Social Contract and Other Later Political Writings*. Ed. Victor Gourevitch. Cambridge UK: Cambridge University Press 1997.

Recebido 11/2022 | Aprovado 02/2023

Living Honestly and Killing Honorably. *Ehre* and *Ehrbarkeit* in the *Metaphysics of Morals*

[Vivendo Honestamente e Matando Honrosamente. Ehre e Ehrbarkeit na Metafísica dos Costumes]

Alessandro Pinzani[1]

Universidade Federal de Santa Catarina (Florianópolis, Brasil/CNPq)

DOI: 10.5380/sk.v21i1.91537

Abstract

This paper aims to call attention to some interpretative questions concerning a semantic constellation running through the entire *Metaphysics of Morals*. This semantic constellation has to do with the German root *ehr-* and refers to such concepts as *ehrlich*, *Ehre*, *Ehrbarkeit*, and *Ehrlichkeit*, which – although they have different meanings – share a connection to the central idea of the coincidence of honesty and honor, as shown by the Latin terms Kant often uses in the same context (*honestas*, *honeste*, etc.). It discusses how Kant uses the concepts of honesty and honor. It tries to answer the question of whether, according to Kant, it is acceptable to die for honor's sake by analyzing some passages from the *Metaphysics of Morals*. Based on this analysis, it finally presents some conclusions on the concepts of dignity and agency and on their reciprocal relationship.
Keywords: Kant; Metaphysics of Morals; Honor; Honesty; Dignity; Agency.

Resumo

Este artigo tem como objetivo chamar a atenção para algumas questões interpretativas que surgem em relação a uma constelação semântica que perpassa toda a *Metafísica dos Costumes*. Essa constelação semântica está relacionada à raiz alemã *ehr-* e refere-se a conceitos como *ehrlich*, *Ehre*, *Ehrbarkeit* e *Ehrlichkeit*, que - embora tenham significados diferentes - compartilham uma conexão com a ideia central da coincidência entre honestidade e honra, como demonstrado pelos termos latinos que Kant frequentemente usa no mesmo contexto (*honestas*, *honeste*, etc.). Discute o modo em que Kant utiliza os conceitos de honestidade e honra. Tenta responder à questão se, na visão de Kant, é lícito morrer por honra, analisando algumas passagens da *Metafísica dos Costumes*. Com base nessa análise, apresenta, finalmente, algumas conclusões sobre os conceitos de dignidade e agência e sua relação recíproca..
Palavras-chave: Kant; Metafísica dos Costumes; honra; honestidade; dignidade; agência.

1 alepinzani@gmail.com. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. I would like to thank Joel Klein, Thomas Mertens, Jean-Christophe Merle, Günter Zöller, and all the colleagues with whom I have discussed this paper. Thanks to Kim Butson for revising the text.

FALSTAFF: What is honour? A word.
What is in that word "honour"?
What is that "honour"? Air.
(W. Shakespeare, Henry IV, Part I, Act V, Scene 1)

This paper aims to call attention to some interpretative questions that arise in connection to a semantic constellation running through the entire *Metaphysics of Morals*, that is, through both its parts: the "Doctrine of Right" and "Doctrine of Virtue". This semantic constellation has to do with the German root *ehr-* and refers to such concepts as *ehrlich*, *Ehre*, *Ehrbarkeit*, and *Ehrlichkeit*, which – although they have different meanings – share a connection to the central idea of the coincidence of honesty and honor, as shown by the Latin terms Kant often uses in the same context (*honestas*, *honeste*, etc.). In Kant's time *honor* and *honestas*, *Ehre* and *Ehrbarkeit* were no secondary concepts in the moral discourse (and in society, as we shall see), and he uses them according to a long tradition that stems from Cicero[2], on the one side, and on the usage of this semantic constellation in the German language, on the other.[3] In both cases, the concepts of honor and honesty are connected to individuals who act in conformity with specific moral codes shared by their society. We shall see that Kant accepts with some restraint this view and relates both concepts to conformity to metaphysical moral principles such as those exposed in the *Metaphysics of Morals*.

1. Honesty in the Metaphysics of Morals

I shall first consider the concept of honesty. In the *Doctrine of Right* it is introduced as *rechtliche Ehrbarkeit* (*honestas iuridica*) in a prominent passage of the "Division of the Doctrine

2 Cicero's *De officiis* was highly influential for the formation of the moral vocabulary of modernity and inspired Kant both directly in the translation of Christian Garve and indirectly through other authors (like Thomasius). Garve, who in 1783 published the translation and commentary of *De officiis* that was known to Kant, translates *honestus* mostly with *moralisch gut* and *vita honesta* with *tugendhaftes Leben* (Cicero, 1783). More recent German translations use *Sittlichkeit* for *honestas* and, according to the context, *sittlich gut*, *sittlich wertvoll*, *sittlich geboten* and *anständig* for *honestum*. Meanwhile, they translate *vita honesta* (*De officiis* III 5-6) both with *sittlich gutes Leben* (see, for instance, Raphael Kühner in Cicero, 1873, p. 184) and with *ehrenhaftes Leben* (see, for instance, Harald Merklin in Cicero 1991, p. 245). No German translation however uses *Ehrbarkeit* for *honestas*. On the other side, the English translator Margaret Atkins stresses the strict connection of the concept of *honestas* with notions of honor (*honor*) and reputation, so that *honestas* can be said to have a public nature. For Cicero, the good man was "in principle *honorable* and would actually be honored by other good men" (Atkins, 1991, p. xlv). For this reason, Atkins translates *honestus* with "honorable" and *honestas* with "honorableness". This is tantamount to establishing an almost full coincidence between honesty and honor. But this perhaps goes too far, since, when referring to the notion of honor in the strict sense, Cicero uses beyond the word *honor* the expression *vir bonus* ("honorable man" or *Ehrenmann*, as the German translators render it) or adjectives such as *nobilis*. Exceptions can be considered the use of the adjective *honesta* when referring to *ars* in order to indicate an honorable profession in *De officiis* I 151 (Cicero, 1991, p. 58) and of the verb *honestare* in *De officiis* II 21 meaning "to promote someone's honor" (Cicero, 1991, p. 70). In the case of *De officiis* II 27, when Cicero says of Sulla that a "dishonorable victory succeeded an honorable cause" [*secuta est honestam causam non honesta victoria*] (Cicero, 1991, p.72 f.), the lack of honorableness follows from the moral unacceptability, and accordingly honor from morality.

3 Among the ancient Germans, the word *Ehre* (in ancient High German: *ēra*) indicated the respect enjoyed by free men and/or their families and clan. It did not depend on richness or birth, but rather on military courage and loyalty. While for the following centuries honor was connected primarily to aristocracy and to the chivalric code, eventually becoming a central concept of the baroque *Hofgesellschaft*, it maintained also a moral meaning. This is expressed particularly through the word *Ehrbarkeit* (that will assume a central role in Kant) and is clearly present in its use by Luther as well as in the common use of the adjective *ehrbar* as synonymous with the Latin adjective *honestus*. Particularly in pietistic circles (such as the one in which Kant was raised), the concept of *Ehrbarkeit* was strictly connected to the individual's behavior toward God, other individuals, and oneself. In the moral works by Thomasius, which were vital in Kant's academic education, it related to Latin notions such as *honestum*, *justum*, and *decorum*: in order to maintain honor, man has above all to fulfill duties to himself by remaining honest. This distinguishes true honor from false and vain honor, which is attached to external things such as richness and birth. (Zunkel, 1975, p. 24 f.). We shall see that Kant maintains the idea that honor is connected to honesty and, therefore, to our way of acting, to agency. On this point see Zunkel, 1975 and the brothers Grimm German Dictionary (DWB).

of Right" (06: 236 f.), namely in connection with the first pseudo-Ulpianian principle "*honeste vive*" (which grounds a quite puzzling "internal" duty of right; see Pinzani, 2005). Interestingly, Mary Gregor translates the term with "rightful *honor*" (Kant, 1996, p. 392), although "legal honesty" could be more appropriate. It consists "in asserting one's worth as a human being in relation to others" and corresponds to the "duty expressed by the saying: 'Do not make yourself a mere means for others but be at the same time and end for them'," which is an "obligation from the right of humanity in our own person".

There are at least two possible readings of this duty, and they are not mutually exclusive. On the contrary, they represent two different ways of "asserting one's worth as a human being in relation to others". The first concerns the prohibition of selling oneself into slavery. This might sound outmoded or archaic, but it was a topic widely discussed by political and ethical writers in Kant's time and earlier in modernity – from Grotius to Rousseau. Kant formulates explicitly the prohibition of self-slavery twice in the *Doctrine of Right*: in § 30 (06: 283) and in General Remark D (06: 330). In this sense, the principle "*honeste vive*" corresponds to what we may call 'the internally mine and yours' as opposed to the "externally mine and yours" that Kant analyses in the following section on "Private Right" (06: 245 ff.). In other words, this principle expresses a duty we have toward ourselves and relates to the idea of self-ownership. We shall see below how slavery is connected also to the concept of honor.

At the same time – so on the second reading – this duty has to do with an internal condition for the existence of Right[4] generally, since only free individuals can enter a juridical relationship with others. While speaking of the "only one innate right", namely external freedom, Kant mentions a "human being's quality of being *his own master* (*sui iuris*), as well as being a human being *beyond reproach* (*iusti*)" (06: 237 f.), with a formulation that alludes once more to the relation between "*honeste vive*" and not being submitted to the arbitrary will of others. An individual who is his own master respects the right of humanity in his person and does not make himself a mere means for others, thereby opening the possibility for the creation of legal relations with other individuals on the basis of legal equality. As one can notice from the position of the "*honeste vive*" in relation to the other two pseudo-Ulpianian principles, particularly to the third ("*suum cuique tribue*", which, in Kant's reading, first refers to the creation of civil society), the obligation of legal honesty can be satisfied also on a pre-juridical level, namely, by affirming one's worth as a possible legal partner. While the State, once created, will care for the protection of the "external mine and yours" and of external freedom, it remains a duty of the individual to care for his own worth and for his 'internal mine'.[5]

This reading will be reinforced if we look at a passage of the *Doctrine of Virtue* in which Kant speaks of the *honestas interna*, which represents the ethical counterpart of *honestas iuridica*. Kant defines this virtue, which he calls "*Ehrliebe*", "love of honor" (!), in negative terms, that is, in the form of a strict prohibition – of a negative perfect duty. This corresponds to the "prohibition against depriving himself of the prerogative of a moral being, that of acting in accordance with principles, that is, inner freedom, and so making himself a plaything of the mere inclinations and hence a thing" (06: 420). As in the case of *honestas iuridica*, *honestas interna* refers primarily to the safeguarding of one's freedom (in this case of internal freedom, since we are in the realm of ethics, not in the realm of Right and external freedom).

According to these two definitions, honesty has to do with maintaining one's freedom (external or internal) and is strictly connected to honor, since it is defined as "love of honor". There is, therefore, a direct connection between *Ehrbarkeit* and *Ehre*, honesty and honor. But, in the next section, we shall see that this connection is not always univocal.

4 I capitalize Right when referring to the main object of the *Rechtslehre* or "objective right", as one says in German, as opposed to right as "subjective right" or claim.

5 Once again, I allow myself to refer to Pinzani, 2005.

2. Honor in the *Doctrine of Right*

Ehre (honor) appears in five different loci within the *Doctrine of Right*. In four cases it is mentioned directly and in one only indirectly. In none of them does Kant offer a clear definition of the concept, though it is possible to understand its meaning in at least three of these loci.

It first appears in the Introduction to the *Metaphysics of Morals*, in which Kant mentions several natural drives: food, sex, rest, movement, and honor (06: 215). While food, sex, rest, and movement do not need to be defined, the same cannot be said of honor, which is a concept whose content is not so obvious, and whose natural character may be called into question. Here, however, Kant is following a tradition of listing human drives or passions that starts from Augustine through Dante and Machiavelli to Hobbes.[6] Furthermore, it is not the first time that our philosopher mentions the drive for honor in his works. In the fourth proposition of the *Idea for a Universal History* (1784) Kant lists "desire of honor [*Ehrsucht*], power or property" as important social motives that drive man "to seek status among his fellows, whom he cannot bear yet cannot bear to leave" (08: 21, here quoted from Kant, 1991, p. 44). In the coeval *Lectures on Anthropology* collected by Christoph Mrongovius (1784/85), Kant introduces a similar triad of "passions," namely "vainglory [*Ehrsucht*], mania for dominance, and greed" (25: 1356, here quoted from Kant, 2012, p. 456), through which humans strive to influence the opinions, fears and interests of other human beings. Kant remarks that "vainglory is not love of honor [*Ehrliebe*]." The latter is

> based on an immediate worth, [...] arises out of modesty and is frank," while the "vainglory is violent, hypocritical" and is based on arrogance, "for it requires others to estimate their worth as less than our own, that is, requires them to be base. [...] Vainglory thus insults and is most hated and resisted. It only wants to get outward respect (25: 1356f.; Kant, 2012, p. 457).

Finally, in the *Religionsschrift* (1793), Kant mentions a slightly different version of this triad, substituting envy [*Neid*] for vainglory [*Ehrsucht*] (06: 94). Therefore, the appearance of *Ehre* in the "Introduction" to the *Metaphysics of Morals* with no further definition can be seen as a reference to something that was commonplace in the philosophical tradition and to other loci in Kant's oeuvre.

In the second locus, the word does not appear directly. I am referring to § 35 of the *Doctrine of Right*, in which Kant speaks of leaving a good reputation after one's death (cf. Mertens, 2019). It is a remarkable passage because Kant says that a good reputation is "an innate external belonging" ("*ein angeborenes äußeres [...] Mein und Dein*"), whose peculiar nature is "that I can and must abstract from whether he ceases to be entirely at his death or whether he survives as a person" (06: 295). This is insofar a puzzling definition, as Kant himself claims that in order to be an external mine something must be an object [*Gegenstand*] that is "distinct from me". Such external objects can be of only three kinds: "1) a (corporeal) *thing* external to me; 2) another's *choice* to perform a specific deed (*praestatio*); 3) another's *status in* relation to me" (06: 247). A good reputation, however, does not belong to any of these three kinds of objects. Furthermore, Kant claims in § 10 that "something external is originally mine which is mine without any act that establishes a right to it," only to remark two lines later: "Nothing external is originally mine, but it can indeed be *acquired* originally, that is, without being derived from what is another's" (06: 258; Kant, 1996, p. 411). So, apparently nothing, including a good reputation, could be

6 Augustine had spoken of three main sins: desire for money and property, desire for power and desire for sexual pleasure (cf. König, 1992, p. 27 f.); Dante encounters, at the beginning of his journey through the afterworld, three beasts symbolizing luxury, envy and greed (Alighieri, 1979, p. 8 ff.); while Machiavelli slightly modified this unholy trinity of negative drives, identifying them with avarice, ambition and an Augustinian desire for "stealing women" (Machiavelli, 1976). In *Leviathan* XIII Hobbes mentions notably three main drives of human action: competition, diffidence and glory (Hobbes, 1996, p. 87 f.), attributing to the latter a preeminent role, since it is the basis for that war of opinions which always threatens peace even after the State has been established.

defined as "an innate external" mine.

But how can Kant contradict himself so blatantly? Or does he really contradict himself at all? The answer lies in an adverb: the good name acquired "by an irreproachable life" is a "(*negatively*) good name" (06: 295; Kant, 1996, p. 442, my emphasis). This means that it is a good name only insofar as it is not besmirched by wrong and unmoral deed – which means in turn that it was originally unblemished. This corresponds to Kant's claim that our one innate right, freedom, implies the quality of "being a human being beyond reproach" (06: 238; Kant, 1996, p. 394). Therefore, we all are born with an untarnished good name (which therefore constitutes an innate belonging) and we may keep it immaculate by living irreproachably. The connection between honor and honesty becomes evident here: by living honestly, I can maintain my innate honor. And by doing so, I acquire the right to see my honor vindicated against slander even after my death, since property is not tied to place and time, as Kant repeatedly points out (06: 245, 247, 249, 254, 296 Fn). This reading may not solve the difficulty of the idea of an innate *external* mine[7], but at least it makes sense of an innate rightful mine, whose existence imposes on others the duty to respect it and creates a corresponding right on my side.

The last three circumstances in which *Ehre* appears are to be found in General Remark E, where the philosopher is discussing crime and criminal law in general, more particularly the death penalty. The first mention of honor in this context is a minor one. If a person offends someone's love of honor, the corresponding punishment might include some form of public humiliation such as being forced to kiss the hand of the offended even when "he is of a lower class" (06: 332). Honor is here evidently used as a synonym for honorability, much in the sense of the good reputation discussed in § 35. In this sense, it is something that might be owned also by persons "of a lower class." This is not the case with honor as defined in the fourth locus, in which it is attributed only to persons of a superior moral nature. The locus is connected to the discussion of Balmerino's case and deserves special attention because it poses relevant questions, both from an exegetical and a theoretical perspective.

In 1745, Arthur Elphinstone, 4th Lord Balmerino, and other Scottish noblemen rebelled against the crown by backing Charles Edward Stuart ("Bonnie Prince Charlie") in his attempt to claim the throne. For this act of rebellion, they were sentenced to death. Kant imagines a case in which the judge would leave the plotters "free to make the choice between death and convict labor." Kant is sure that "the man of honor would choose death and the scoundrel forced labor." His explanation is the following: "This comes along with the nature of the human mind; for the man of honor is acquainted with something that he values even more highly than human life, namely *honor*, while the scoundrel considers it better to live in shame than not to live at all" (06: 333 f.; Kant, 1996, p. 474f.). Apparently, according to this example, there are two kinds of men, at least among plotters: men of honor and scoundrels. Two sets of important questions arise here. First: how are the former to be distinguished from the latter? What transforms an individual into a man of honor? How is honor defined in this context? Second: may we value honor more than life, as the man of honor is supposed to do? Is this preference in accordance with the moral law that requires us to respect humanity in our person and forbids us to take our own life? The first questions have interpretative importance since they refer to something internal to Kant's theory. The second have a wider philosophical relevance that goes beyond the mere Kantian exegesis. In the next two sections, we answer both questions.

3. Who is a Man of Honor?

We begin with the first questions. It is not very clear how Kant comes to distinguish between men of honor and scoundrels in Balmerino's rebellion. While discussing this case, he

7 Mertens (2019, p. 518f.) suggests a plausible answer to this question.

states an important premise, supposing that some of the rebels "believed that by their uprising they were only performing a duty they owed to the House of Stuart, while others were out for their private interests" (06: 333; Kant, 1996, p. 474). We might think that the men of honor were the former, while the latter were the scoundrels: while the scoundrels were up to no good, the men of honor were acting in good faith or *bona fide* (a Latin term that Kant translates always with "*ehrlich*"[8], connecting it therefore with *Ehre* and/or *Ehrbarkeit*).[9]

According to Kant, "the man of honor is undeniably less deserving of punishment" than the scoundrel (06: 334; Kant, 1996, p. 475). That does not mean he does not deserve capital punishment, but just that this is the only kind of punishment proportionate to his sensibility (since an honorable death is more important to him than an abject life as a slave) – convict labor would represent an exceedingly severe punishment for him. So, one could be sentenced to death and maintain one's honor, *if one was already a man of honor*, that is, if one's criminal actions were undertaken in good faith or *bona fide* (in Kant's translation: *ehrlich*). This, however, in no way represents a mitigating circumstance and does not save one from punishment, because – as we all know – Kant is adamant that a man who has committed murder must die (06: 333; Kant, 1996, p. 474).[10] He is very strict on this, to the point of practically excluding the possibility of grace (06: 337; Kant, 1996, p. 477).[11] Therefore, that Balmerino and some of his accomplices were "men of honor" because they acted in good faith does not represent a justification for their deeds, nor can it be used to advocate an exemption from the death penalty. Accordingly, the rebels *ought* to be put to death even if they remain men of honor.

Does this affect their *honestas iuridica*, their rightful honesty? Since they have committed a crime, they surely lose the right to be held "beyond reproach" (06: 238), even if they were acting *bona fide*. As we will see below, criminals lose not only their honesty and their external freedom (and in some cases, their lives) but also their personality and dignity. The connection between honor and honesty seems, therefore, to be lost, in the case of Balmerino and his accomplices: they might remain men of honor, but they can no longer be seen as honest men.

4. Can We Renounce Life for Honor's Sake?

To turn to the second questions: is it legitimate for a man of honor to prefer death to convict labor and, therefore, to a life "in shame"? Does honor justify sacrificing one's life? Should we not strive to preserve our life on every occasion? Do we not have a perfect duty to ourselves regarding self-preservation?

While discussing this duty in the *Doctrine of Virtue*, Kant seems to allow for exceptions. In the corresponding casuistic questions (06: 423f.; Kant, 1996, p. 548) he mentions some cases, in which it is *apparently* permissible to kill oneself. However, it is difficult to state whether he really

8 Kant uses several times the adjective *ehrlich* in connection to ownership or to the person of the owner (for instance 06: 292, 300 ff., 364). In all these cases the term translates explicitly the Latin expression "*bona fide*".

9 It is very unlikely that Kant uses the expression "*ehrlicher Mann*" ("man of honor") as synonymous with "*Edelmann*" ("aristocrat"), as was common in his time. In all his political writings Kant is coherent in rejecting hereditary (blood) aristocracy as a form of unjustified, irrational privilege.

10 On Kant's view of the death penalty see Ataner 2006 and Merle 2009, 44-71.

11 The sovereign may grant clemency only in the case of a wrong done to himself, not in the case of crimes against a subject. Murder is therefore excluded from the crimes for which clemency is possible, since the sovereign who could grant it as the victim of this crime would be dead.
In a much quoted, quite infamous passage, he writes: "Even if a civil society were to be dissolved by the consent of all its members (e.g., if a people inhabiting an island decided to separate and disperse throughout the world), the last murderer remaining in prison would first have to be executed, so that each has done to him what his deeds deserve and blood guilt does not cling to the people for not having insisted upon this punishment" (06: 333). This old Testamentary wording may surprise the reader, but it fits to the conception of punishment as retribution or retaliation (*ius talionis*) that Kant defended (06: 332).

believes that the strict prohibition of suicide can be violated in any case at all. First, he does not affirm anything positive about this point: he just formulates casuistic questions and does not provide any answers. Second, the cases mentioned differ greatly from one another. In some cases, death is certain anyway and the choice for suicide appears therefore quite unproblematic: this is the case of Seneca (who killed himself in anticipation of "the unjust death sentence" handed down by Nero) and of the hydrophobic man (since no remedy for or vaccination against rabies was available at that time). In other cases, death is not certain, but can be seen as a way of avoiding greater harm to others: this is the case of Curtius (who sacrifices himself to save his country), of Frederick the Great (the "great king" who carried a fast-acting poison with him to spare any harm to his state if he were captured in battle) and, once again, of the hydrophobic man (who does not want to harm others). In one case, death is just a possible outcome of an action (vaccination), whose aim is exactly to protect one's life and to reduce the risk of death (in this case the infection through deadly pathogenic agents like smallpox); therefore, it is not a proper case of suicide.

The case of Balmerino's putative choice for death over convict labor does not fit completely in any of these possibilities.[12] While Curtius, Frederick, and the hydrophobic man chose death in order not to harm others, Balmerino did so in order to maintain his honor and not live in shame. How can this egoistic motivation be put on the same level as the altruistic motives of the other examples?

Two answers are possible to this question. The first appeals to Kant's recognition of the central relevance of honor in his society and can be backed by the fifth locus of the *Doctrine of Right* in which Kant mentions honor. The second refers once again to the problematic connection between honor and honesty. We shall consider both answers in the following sections by analyzing the fifth appearance of honor in the *Doctrine of Right*.

5. Killing for Honor's Sake

In the same General Remark E in which Kant discusses Balmerino's case, we find the fifth locus, in which honor plays a role. In it, Kant mentions "two crimes deserving of death, with regard to which", however, "it still remains doubtful whether legislation is also authorized to impose the death penalty" (06: 335 f.; Kant, 1996, p. 476). *Nota bene*: what is presented as doubtful here is whether Right has the authority (*Befügnis*) to impose capital punishment, not whether it might be effective in threatening the criminal with the death penalty, as is the case with *Notrecht*, with that "right of necessity" which does not even belong to Right strictly (06: 235 f.).[13] Furthermore, it is not clear whether Kant himself considers this to be a doubtful case, for he

12 Of course, he was sentenced to death like Seneca, but justly and through a rightful sentence, contrarily to the Roman philosopher. But this is not relevant here.

13 In that case, Kant uses Carneades' classical example of two survivors of a shipwreck who try to save themselves by grasping a plank on which there is space only for one of them. According to Kant, it is impossible to legally forbid one of them to push the other into the water since, if he doesn't do this, he will die. The threat of a penalty, even of the most severe possible (i.e., the death penalty, which is what should restrain him from killing someone under normal circumstances) would have no effect at all, so this case cannot possibly be regulated by Right since the latter would have no power at all over the individual whose actions it should punish. This is why, in Kant's vision, it is a very special case that does not even belong properly in the *Doctrine of Right*: the involved persons find themselves in a situation of life or death, in which Right is impotent; it is as if Right were suspended and they had fallen back into the state of nature, in which there is no law restraining them and no worse punishment other than violent death by the hand of others. That does not mean, however, that I am not doing anything wrong by pushing someone off the plank to save myself. The fact that I cannot be restrained by any legal threat from doing so does not imply that what I am doing is not an *Unrecht*, a wrongful act. It is not by chance that Kant speaks here of an "alleged right" (Kant, 1996, p. 391). He distinguishes between the objective and the subjective dimension of the claim that I am authorized to kill someone to save my own life. The subjective dimension refers to the fact that no legal threat can stop me, since the most powerful one, that is, the promise of the death penalty, threatens me precisely with what I shall meet anyway if I yield to it: I will drown if I do not kill the other. Furthermore, while the death penalty is uncertain (who

uses the ambiguous formulation "it remains doubtful" (in German: "noch zweifelhaft bleibt"). This expression can just represent a statement on a matter of fact (someone has questioned the authority of the state) that does not necessarily correspond to Kant's own opinion (he might think that the state absolutely has this authority). Alternatively, it can constitute a way of introducing the analysis of the following crimes, meaning more or less: "we still have to consider two cases that are *prima facie* doubtful". The impression that Kant himself does not have any doubts about the real criminal nature of the two cases might be reinforced by his use of the verb *verleiten*, whose distinctly negative, condemning touch gets lost in Mary Gregor's English version – which uses the morally neutral verb "to lead". But an even stronger argument in favor of Kant's lack of doubts is the open condemnation of killing for honor's sake that he formulates at the end of the paragraph.

Let us consider the first crime, which is quite shocking and often provokes uneasiness among Kantian scholars. Kant claims that a woman killing her child *apparently* [*so scheint es*] does not deserve the death penalty and is not to be considered a murderer at all if the baby resulted from an illegitimate relationship: this would be a mere act of killing [*Tötung*], not a murder [*Mord*] (06: 336).[14] The justification for this is that the child who is born outside the protection of marriage is *eo ipso* outside the law and, therefore, has no legal protection (he is not a citizen, not even a passive one). Kant speaks here of the baby as if he were a thing: "It has, as it were, stolen into the commonwealth (like contraband merchandise), so that the commonwealth can ignore its existence (since it was not right that it should have come to exist in this way), and can therefore ignore its annihilation" (06: 336; Kant, 1996, p. 477).

If this was really Kant's position, then it would mean that nobody had a right to life for the mere fact of being born, for existing as a member of the species. Children born within the context of a legal marriage are protected by the Law and their parents are not allowed to kill them (or sell them or handle them as things – cf. 06: 280 ff.), contrary to what happens with children born outside of marriage. That is, children receive full legal protection only when they are born inside a *legally* defined and sanctioned situation (i.e., marriage). Kant does not say, however, that it is legitimate to kill a baby born outside a legal marriage, but only that the *mother* may do so without becoming a murderer in the eyes of the Law. He does not even claim that the community has the right to ignore the killing, he just observes that it can do this (the use of the verb "*kann*" instead of "*darf*" should be noted here). Apparently, Kant is just registering a fact: people do not consider this act a proper murder and prefer to ignore it. But how is this possible? How can the killing of a baby born outside a legal marriage not be considered murder? Kant answers this question by recurring to the notion of honor. According to the socially dominant view, a woman who suffers the loss of her honor has the right to eliminate the cause of the loss. By killing the illegitimate child, the mother eliminates the cause of her disgrace (*Schande*).

---

will denounce me if I am the only survivor?), death by drowning is certain. Therefore, subjectively I see myself as authorized to kill the other. It is *me* who authorizes *myself*. The objective dimension refers to Right [*Recht*] itself, and Right cannot accept me killing an innocent to save my own life. Therefore, there is no right such as the right of necessity (Merle, 2021).

14 On the relevance of this topic in 18th century Germany see Kord, 2009, p. 121 ff. According to her, "seduction and, during the second half of the century, infanticide can fairly be called *the* most popular themes in eighteenth-century German literature authored by men" (Kord, 2009, p. 121). While in the literature of previous centuries, seduction brought about mostly humorous situations and served as material for comedies, in the 18th century a significant "conceptual shift" occurred: the inexperienced girl's consent to premarital sex, "which was, until the eighteenth century, easily repaired through marriage, metamorphoses, in the eighteenth century, into a loss of 'honour' and attending sense of shame that inevitably results in death" (Kord, 2009, p. 122). As the author points out, "the topic of infanticide attracted the attention of philosophers and legal reformers as well as literary authors", many of whom, interestingly, "were also or had been students of law" (Ibid..). There is, however, a strange anachronism in Kant's depiction of infanticide. While he acknowledges doubts concerning the state's authority to punish infanticidal mothers, the legal reality of his time was that all legal orders foresaw the death penalty for this crime. It is not by chance, that all the heroines of the mentioned fictional works met this tragic fate (e.g., Evchen in Heinrich Leopold Wagner's *Die Kindermörderin*, Marie in Johann Michael Reinhold Lenz's *Zerbin* and, most famously, Gretchen in Goethe's *Faust*).

This is a very interesting case, since the child, contrarily to the criminal (whose case we shall consider below), has done nothing to forfeit his dignity or humanity: it is rather the mother who acted against custom by conceiving outside marriage. Yet, the child has done nothing to preserve his dignity either, since he cannot act as a moral agent. He seems to inhabit a sort of moral limbo: he is already a member of the species, but he is not yet a person with humanity or dignity. One could even doubt whether he can be considered fully alive given Kant's definition of life in the Preface to the second *Critique*: "Life is the faculty of a being to act in accordance with laws of the faculty of desire" (05: 9, footnote; Kant, 1996, p. 144). Does a newborn baby have such a faculty? He does, but only potentially. The illegitimate baby is only *potentially* a living being and only *potentially* a moral agent. But *actually*, he represents a loss of honor for the mother – that was what society in Kant's times thought, anyway – and honor must be taken seriously by the Law, at least for the time being.

Kant surprises us here with a quasi-Hegelian remark: In another, more civilized and less "barbarian" time, having an illegitimate child would no longer be a reason for disgrace, so killing him would be considered plain murder. But, until that time, the categorical imperative of penal justice (06: 336) must yield to the mores, morality must yield to costumes. For the same reason, Kant claims that there is another case in which killing someone cannot be considered murder: a duel, which represents the second exception to the rule of *ius talionis* that characterizes Kant's interpretation of criminal law. If my honor has been unduly violated, I may kill the offender in a duel.[15] Again, this may be considered a barbarian and uncivilized thing to do, but as long as we assign such preeminence to honor, we have to accept that people kill for honor's sake and that Right has no authority to demand death for this kind of killing. If penal justice punished these actions with death, that would be tantamount to declaring that the concept of honor ("which is *here* no illusion", as Kant remarks – italics mine) "counts for nothing" (06: 336; 1996, 477). According to Kant, the only way out of this "quandary" is that the concept of honor is transformed in such a way that what counts at present as an offense to it can then be considered less important and, therefore, is no longer a justification for "unlawful killing."

The question about the legitimacy of choosing death in honor over living in shame finds here its first answer: "Here" (i.e., in this society and in this specific historical time) honor is no illusion and may count for something. Therefore, it is justifiable to kill and to die for honor's sake, and one questions even the state's authority to intervene with the death penalty against honor killers. But, for Kant, this situation is historically contingent and does not deserve to be preserved, for it is an expression of a time, whose legislation is still "barbarous and underdeveloped" (06: 337; Kant, 1996, p. 477). Therefore, the first answer, according to which it is only because society gives an undeserved relevance to honor that we consider it acceptable for Balmerino to choose death, is not completely satisfactory (and somehow trivial). The second answer is not as directly connected to Kant's treatment of honor as the first one, but it is far more interesting since it helps us to grasp better not only the connection between honor and honesty, but also the link between honesty, dignity, and freedom. To start discussing it, we should first consider how honor is treated in the *Doctrine of Virtue*.

6. Honor in the *Doctrine of Virtue*

Compared to the cases discussed in the *Doctrine of Right*, the references to honor in the *Doctrine of Virtue* are quite harmless. The first connects love of honor to internal honesty (*honestas interna*) and to internal freedom, as we have already seen. The second occurs in § 42 and § 43, in which Kant speaks of the vices of arrogance and defamation. Arrogance is defined as a kind of ambition [*Ehrbegierde*, i.e. literally: desire of honor], "in which we demand that others

15 On duels in general and on Kant's view on dueling see Stell, 1979.

think little of themselves in comparison with us". It is to be distinguished from "pride proper [...], which is love of honor [*Ehrliebe*]" (06: 465; Kant, 1996, p. 581).[16] The arrogant seeks honor (he is *ehrsüchtig*, an addict to honor), but at the same time, he treats with contempt the very people from whom he expects recognition. This attitude is not only unjust but is also foolish, more precisely: it shows *Torheit* and *Narrheit*.[17] In the case of defamation, the vice consists in "the intentional *spreading* (*propalatio*) of something that detracts from another's honor" (06: 466; Kant, 1996, p. 582).

In the case of defamation, honor obviously corresponds to the good reputation that was the object of the abovementioned § 35 of the *Doctrine of Right* (06: 295; Kant, 1996, p. 441f.). In the case of ambition [*Ehrbegierde*], it is seen as something we must strive for, and which depends on other people's recognition. Pride proper is defined there as "a concern to yield nothing of one's human dignity in comparison with others" (06: 465; Kant, 1996, p. 581), while at the same time respecting the honor of these others (when this respect is not shown, one is behaving arrogantly). Once again, honor is connected to dignity, and this is understood as something belonging to the individual as a human being ("human dignity"). However, individuals may lose not only their honor and (rightful) honesty but precisely their human dignity and even their personality, becoming a thing.

7. Losing One's Dignity and Personality

In General Remark D, Kant states that "no human being in a state can be without any dignity, since he at least has the dignity of a citizen.[18] The exception is someone who has lost it by his own crime" (06: 329; Kant, 1996, p. 471). The criminal loses his dignity as a citizen, and therefore presumably his honor. However, it seems that he also loses his dignity as a human being, because, as Kant claims, "though he is kept alive, he is made a mere tool of another's choice [*Willkür*] (either of the state or of another citizen)". This means that, apparently, he may be considered as a mere means and no longer as an end in himself; he can be reduced to slavery and as such 'given' by the state to a private citizen, who evidently may use him as a tool for work – as still happens in some countries, in which inmates are forced to work for the state (like in China) or for private companies (like in some U.S. states) without receiving a wage. The passage continues in the following terms:

> Whoever is another's tool (which he can become only by a verdict and right) [which excludes the possibility of self-slavery: one can be condemned to slavery, but may not make himself a slave] is a *bondsman* [*Leibeigener*] (*servus in sensu stricto*) and is the *property* (*dominium*) of another, who is accordingly not merely his *master* (*herus*) but also his *owner* (*dominus*) and can therefore alienate him as a thing, use him as he pleases (only not for shameful purposes) and *dispose of his powers*, though not of his life and members.

As Kant puts it quite bluntly, "he ceases to be a person", not simply to be a citizen (06: 330; Kant, 1996, p. 471f.). He is no longer his own master (therefore he is no longer an honorable man) and has become a thing, the property of another, a mere tool in his owner's hands. Can we say that he still has dignity as a human being? What about the right of humanity in his person? Does the categorical imperative still apply to him? Apparently, it does, but only with regard to the prohibition to use him for shameful purposes – which however could be seen more as an obligation of the owner toward himself, that is, as a perfect duty to oneself, as

16 Cf. the abovementioned passage from the *Lectures on Anthropology* collected by Mrongovius.

17 The English translator calls our attention to the distinction Kant made in *Anthropology from a Pragmatic Point of View* (7: 210) between being "foolish" [*Tor*] and "a conceited ass" [*Narr*].

18 On the different meanings of "dignity" in Kant and its relation to Right see Von der Pfordten, 2009 and Sensen, 2011.

expressed in § 7 of the *Doctrine of Virtue* (06: 424 ff.; Kant, 1996, p. 548ff.) – and with regard to the integrity of life and limbs. In any other regard, he is just a thing, and his owner may use him (or better: it) as he pleases.

This is not the only instance in which an individual becomes dependent on another. Kant observes here that "no one can bind himself to this kind of dependence" (once again we observe his criticism of the idea of voluntary self-slavery) and that those who put themselves under obligation to another person by contract become their subjects, not their bondsmen (06: 330; Kant, 1996, p. 472). However, according to a passage in which Kant considers this question, namely § 30, it seems that by this contract they almost completely lose their external freedom. In this case, he speaks about servants who are hired by the head of a household, who "are included in what *belongs* to [him] and, as far as the form (the way of his being in *possession*) is concerned, they are his by a right that is like a right to *a thing*" (06: 283; Kant, 1996, p. 431, my emphasis). In § 29 (06: 282; Kant, 1996,p. 430), while referring to the relation of the father to his children, Kant had discussed "a right to a person *akin to a right to a thing*" (*ein auf dingliche Art persönliches Recht*), which allows parents to take control of their children and to impound them as things "like domestic animals that have gone astray". The idea of fugitive animals seems to have inspired Kant also when he writes that, if servants run away from their master, "he can bring them back in his control by his unilateral choice", that is, without asking permission from a tribunal. He can run after them, seize them and bring them back as if they were runaway horses.

Nevertheless, children and servants maintain their dignity and cannot be seen fully as things. The parents and the master may not dispose of them as they please. The right the father viz. master holds over them is still "a right to a *person*", even if it is "akin to a right to a thing". This is a completely different case from that of "the right of ownership with regard to someone who has forfeited his personality by a crime" (06: 283; Kant, 1996, p. 431). As becomes clear from Kant's wording, the criminal forfeits his personality [*Persönlichkeit*], not merely his dignity as a citizen. What is at stake, in this case, is more than simply honor or legal honesty: it is personality itself. How can we conciliate this view with the one Kant expressed in his ethical works, in which human beings have a distinct dignity as autonomous moral lawgivers? Does the criminal lose also *this* dignity? Does he cease to be an autonomous member of the Kingdom of Ends?

Metaphysics of Morals does not appear to hold an answer to this question. One reason for this might be that in works such as the *Groundwork* or the second *Critique* individuals are seen primarily as rational beings, as *homines noumena* moved by reason and autonomous authors of the moral law, more than as *homines phenomena* moved by passions and renitent subjects of the moral law. In both parts of the *Metaphysics of Morals*, on the contrary, Kant refers to man as *homo phenomenon*, as someone who is trying to satisfy his needs and desires (in the *Doctrine of Right*) and to organize his life according to the moral law (in the *Doctrine of Virtue*). While the first issue involves dealing with the needs and desires of others, so that the main task is to coordinate one's free choice [*Willkür*] with the free choice of others (according to Kant's definition of Right), the second issue leads to a perpetual struggle to control one's motives and passions, that is, to develop virtues that help to make the moral law into the habitual maxim of one's behavior. In both cases, honesty has to do with affirming one's worth as an agent: a legal and a moral one. (*Honestas iuridica* and *honestas interna* respectively refer to these kinds of agency.) According to Kant's own systematic and architectonic division in the "Introduction to the Metaphysics of Morals," the two aspects should remain separate, but – as we saw – the loss of legal honesty, and therefore of legal personality, may lead to the loss of internal honesty, and therefore of personality in general.

We should make here a distinction, which Kant does not make explicitly, but that he seems to make implicitly. One can be a human being, a *Mensch*, in two different senses: first, as a member of the species; second, as someone who has dignity because of his humanity, of

his *Menschheit*, which is to be understood as the capacity to be externally free (Pinzani, 2018, p. 222). Now, a human being cannot cease to be a member of the species, but he can forfeit his humanity and therefore his dignity. Contrarily to membership of the species, humanity is something we must keep through our deeds. It depends on what we do, that is, on the kind of person we become through our actions. If we do the wrong thing, we may lose it and, by losing it, we may lose our external freedom and its corresponding "authorizations," including the right to be our own masters and to be beyond reproach (06: 238; Kant, 1996, p. 393f.). It is true that our "innate personality" prohibits us becoming "mere tool(s) of another's choice", but if we are found "punishable" for a crime, then this prohibition is no longer valid (06: 331; Kant, 1996, p. 473).

Far from being the object of an absolute, untouchable right grounded in innate dignity, life has to do with our capacity to be legal and moral agents. One could say that dignity and life are something that you should *deserve*. They depend on our acting; therefore, dignity is itself based on agency, exactly as life has value only for being the necessary condition for acting. We can lose our dignity (and our life) because of certain actions, even if we remain rational beings; or, better yet, precisely because we are rational beings able to be held responsible for our actions. According to Kant, we do not have rights as human beings, but rather as rational agents who assume full responsibility for their actions and who must bear the consequences. Agency becomes, therefore, literally a matter of life or death, since under certain circumstances, by acting against the Law, one can forfeit his right to the legal protection of one's life by the State or can even lose his right to be considered a worthy human being, becoming instead a thing that can be possessed and disposed of at its master's pleasure. This is a quite disturbing conclusion for our traditional vision of Kant's practical philosophy. Yet, it opens up an interesting discussion on the right to life and on the right to kill that may extend much further than mere Kantian exegesis.

Bibliographic References

ALIGHIERI, D. *La Divina Commedia: Inferno*. A cura di U. Bosco e G. Reggio. Firenze: Le Monnier, 1979.

ATANER, A. 'Kant on Capital Punishment and Suicide'. In: *Kant-Studien*, 97, 2006, p. 452–482.

ATKINS, E. M. 'Introduction'. In: CICERO, Marcus Tullius. *On Duties*. Cambridge: Cambridge University Press, 1991, p. ix-xxviii.

CICERO, M. T. *Abhandlung über die menschlichen Pflichten*. Translation and commentary by Christian Garve. Breslau: Korn, 1783.

CICERO, M. T. *Von den Pflichten*. Übersetzt von Raphael Kühner. Stuttgart: Hoffmann, 1873.

CICERO, M. T. *Von den Pflichten*. Übersetzt von Harald Merklin. Frankfurt a. M.: Insel, 1991.

CICERO, M. T. *On Duties*. Edited by M. T. Griffin and E. M. Atkins. Cambridge: Cambridge University Press, 1991.

DWB = Deutsches Wörterbuch von Jacob und Wilhelm Grimm. 16 Bde. in 32 Teilbänden. Leipzig 1854-1961. Quellenverzeichnis Leipzig 1971 (http://woerterbuchnetz.de/cgi-bin/

WBNetz/wbgui_py?sigle=DWB&mainmode=)

HOBBES, T. *Leviathan*. Edited by R. Tuck. Cambridge: Cambridge University Press, 1996.

KANT, I. *Political Writings*. Edited by H. S. Reiss. Cambridge: Cambridge University Press, 1991.

KANT, I. *Practical Philosophy*. Translated by M. Gregor, edited by A. Wood. Cambridge: Cambridge University Press, 1996.

KANT, I. *Lectures on Anthropology*. Edited by A. Wood and R. Louden. Cambridge: Cambridge University Press, 2012.

KÖNIG, H. *Zivilisation und Leidenschaften. Die Masse im bürgerlichen Zeitalter*. Reinbek bei Hamburg: Rowolth, 1992.

KORD, S. *Murderesses in German Writing 1720-1860. Heroines of Horror*. Cambridge: Cambridge University Press, 2009.

MACHIAVELLI, N. Il Principe e le Opere Politiche. Milano: Garzanti, 1976.

MERLE, J-C. *German Idealism and the Concept of Punishment* Cambridge: Cambridge University Press, 2009.

MERLE, J-C. Das „zweideutige Recht" („Anhang zur Einleitung in der Rechtslehre"). In: MERLE, J.-C.; FREIIN von VILLIEZ (Hgs.). *Zwischen Rechten und Pflichten – Kants „Metaphysik der Sitten"*. Berlin: de Gruyter, 2021, p. 95-104.

MERTENS, T. '"Bona fama defuncti" in Kant's "Rechtslehre": Some Perspectives'. In: *Kantian Review*, Vol. 24, n° 4, 2019, p. 513-529.

PINZANI, A. 'Der systematische Stellenwert der pseudo-ulpianischen Regel in Kants "Rechtslehre"'. In: *Zeitschrift für philosophische Forschung*. Vol. 59, n°1, 2005, p. 71-94.

PINZANI, A. 'Recht der Menschheit und Menschenrechte'. In: MOSAYEBI, R.(Hg.). *Kant und Menschenrechte*. Berlin: de Gruyter, 2018, p. 217-237.

SENSEN, O. *Kant on Human Dignity*. De Gruyter, Berlin, 2011.

STELL, L. 'Dueling and the Right to Life'. In: *Ethics*, Vol. 90, n°1, 1979, p. 7-26.

VON DER PFORDTEN, D. *Menschenwürde, Recht und Staat bei Kant. Fünf Untersuchungen*. Paderborn: Mentis, 2009.

ZUNKEL, F. 'Ehre, Reputation'. In: BRUNNER, O.; CONZE, W.; KOSELLECK, R. (Hgs.). *Geschichtliche Grundbegriffe*. Band 2. Stuttgart: Klett-Cotta, p. 1975, 1-63.

Recebido 11/2022 | Aprovado 12/2022

Passions and social irrationality

[Paixões e irracionalidade social]

Maria Borges[1]

Universidade Federal de Santa Catarina (Floranópolis, Brasil/CNPq)

DOI: 10.5380/sk.v21i1.91538

Abstract

In this paper, I analyze the concept of passions in Kant, trying to answer the question: are passions good for society or are they the source of evil and irrationality? I begin by showing the danger of passions, if compared with affects. In the *Anthropology*, as well in the *Religion*, Kant claims that passions are the evil we should fight against. More than inclinations, or affects, the evil principle refers to passions. However, in the *Ideia for an Universal history with a cosmopolitan aim*, Kant admits that passions are good for society. They are not the cause of their irrationality, but the possibility of the development of the natural predispositions of human being. I claim that we have two distinct answers in Kant, and I try to understand this ambiguity.

Keywords: passions; irrationality; history; morality; vainglory; greed.

Resumo

Neste artigo, analiso o conceito de paixões em Kant, buscando responder à questão: as paixões são benéficas para a sociedade ou são fonte de mal e irracionalidade? Começo mostrando o perigo das paixões, quando comparadas aos afetos. Na *Antropologia*, assim como na *Religião*, Kant afirma que as paixões são o mal contra o qual nós devemos lutar. Mais do que inclinações ou afetos, o princípio do mal refere-se às paixões. No entanto, na *Ideia de uma história universal de um ponto de vista cosmopolita*, Kant admite que as paixões são benéficas para a sociedade. Elas não são a causa de sua irracionalidade, mas a possibilidade de desenvolvimento das predisposições naturais do ser humano. Eu afirmo que temos duas respostas distintas em Kant, e procuro compreender essa ambiguidade.

Palavras-chave: paixões; irracionalidade; história; moralidade; vanglória; ganância.

1 mariaborges@yahoo.com. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

People think of Kant as having a moral portrait of human being, as someone who mostly acts according to the moral law. If we accept this intrinsic moral and good-hearted representation, we fail in understand essential features of Kantian account. First, that the moral law acts upon the human mind as an imperative, not as a natural way of behaving. Second, that Kant has a normative ethics, he says what we ought to do not what we generally do. Third, that he has a theory about human passions that show they are forces who oppose the moral law.

The danger of passions

In the *Anthropology*, in comparing affects with passion, Kant shows the strong and damaging effects of it (Anth, 7:252).

> Affect works like water that breaks through a dam; passion, like a river that digs deeper and deeper into its bed. Affect works on our health like an apoplectic fit, passion, like consumption or emaciation. Affect is like drunkenness that one sleeps off, although a headache follows afterward; but passion is regarded as a sickness that comes from swallowing poison, or a deformity which requires an inner or an outer physician of the soul, one who nevertheless knows how to prescribe remedies that are the most part radical, but almost always merely palliative (Anth, AA 7: 252).

Affects are less harmful than passions: affects are like drunkenness, while passion is a disease resulting of taking poison, or a deformity which requires a physician. Affect is less permanent than passion: it is like an apoplectic convulsion, while passion is a consumption or atrophy. Besides that, the remedies prescribed by a physician to passions are radical, but palliative. It seems that, while one can get rid of affects or cure them, passions, however, are never really healed.

Although passions can be considered a disease of the mind, they can imitate reason:

> Since the passions can be coupled with the calmest reflection, one can easily see that they must neither be rash like the emotions, nor stormy and transitory; instead, they must take roots gradually and even be able to coexist with reason (Anth, AA 7: 266).

Passions are not transitory and stormy like affects, but they can coexist with reason and deliberation, forming principles and/or maxims upon them.

> A passion is a sensible desire that has become a lasting inclination (e.g., hatred, as opposed to anger). The calm with which one gives oneself up to it permits reflection and allows the mind to form principles upon it and so, if inclination lights upon something contrary to the law, to brood upon it, to get it rooted deeply, and so take up what is evil (as something premeditated) into its maxim. And this evil is then properly evil, that is, true vice" (MS, AA 6:408).

Passion is responsible for forming evil maxims, based on reflection and premeditation. They are not like affects, that are only frailty, inclinations that make us act against moral law in a stormy way.

Passion as the last and strongest moment of desire

What is the place of passion in the faculties of the mind? It is usually considered the last, and strongest, moment of the faculty of desire. According to the Anthropology Mrongovius, the kinds of faculty of desire are propensity, instinct, inclination and passion. (V-Ant/Mensch, AA 25:1339). Propensity is the origin of desire, that "can occur even if the desire is not there yet". As examples, we have the north people's tendency to drink strong drinks. The second is

instinct (*Instinkt*), which consists in a desire without previously knowing the object by which it is satisfied, e.g., a child's instincts desire for milk. The third kind is inclination (*Neigung, inclinatio*), is a habitual desire, and exemplified with the desire to play games or drink. If an inclination is too strong, it becomes a passion (*Leidenschaft, passio animi*), which is the last degree of the faculty of desire. Passion is the "inclination which can hardly, or not at all, be controlled by reason" (Anth, AA 7:251), and "with passion, one is not able to compare the inclinations with the sum of all other ones" (V-Ant/Mensch, AA 25: 1340).

Kant uses inclination both in a broad and in a specific sense; in the first meaning, all empirical incentives can count as inclinations, in the second, inclination is only the third degree of the faculty of desire. Passion is an inclination in the first sense, as an empirical incentive to action. It is stronger than the other inclinations, even the sum of all that, showing its robust and resilient nature.

Natural and Social passions

Kant categorizes passions into natural and social ones. Natural passions are called "burning passions", they are the inclinations for freedom and sex. The social passions are the cold ones: mania for honor (*Ehrsucht*), mania for domination (*Herrschsucht*) and mania for possession (*Habsucht*) (Anth, AA 7:272-275).[2] The passion of freedom is the natural desire not to depend on other people: "whoever is able to be happy only at the option of another person, feels that he is unhappy" (Anth, AA 7:268). It is a natural desire, a desire to keep others far away, and to live "as a wanderer in the wilderness". It is a desire of men in the state of nature, before the civil society is established.

The three main social passions (Ehrsucht, Herrschucht, Habsucht) are related to the three ways of having influence on human beings, And they also have their particular age: "Vainglory/ mania for honor belongs to youth , the mania for dominance to manhood, and the greed to old age". All social passions are a way to have influence on other human beings: "by means of honor, we have influence on the opinions of human beings, by means of authority we have influence on their fear, and by means of money we have influence on their interest" (V-Anth/ Mensch, AA 25: 1356).

The mania for honor, arrogance, and the flatterers

Kant distinguish love of honor (*Ehrliebe*) from vainglory (*Ehrsucht*):

> Vainglory is not love of honor. Love of honor is based on an immediate worth, but vainglory is based on a mediate one, namely, insofar as it is of service for having influence on others in other people's eyes. Love of honor arises out of modesty and is frank, Vainglory is violent, hypocritical. (...) One can thwart the aims of the vainglorious person most easily. One can slight him without even saying anything to him, merely through indifference (AntM, AA 25:1356-57).

The one who has mania for honor is usually also arrogant. The arrogance needs to be flattered in order to be sure of its value, and indifference is for him the worst reaction. That is the reason why the flatterer can conquer and even destroy powerful man, using this special spell towards the vainglorious person: "Flatterers, the yes-men who gladly concede high-sounding talk to important man, nourish this passion that make him weak, and are the ruin of the great

2 Translation of Ehrsucht, Herrsucht, Habsucht: Clélia Martins (Anth, AA, 7:271), ambição, desejo de dominação e cobiça; Robert Louden (Anth, AA, 7: 271.), mania for honor, mania for domination, mania for possession.

and powerful who abandon themselves to this spell" (Anth, AA 7: 273).

A good flattering is a spell, and powerful men can be destroyed by this. One may ask: how powerful people could be ruined by flatterers, since they are probably not naïve persons? Vainglory, as a passion, makes the subject blind. He wants glory and fame so badly, that his vanity makes him an easy catch for people who knows how to seduce. The flatterers use an idealized mirror, making the vainglorious person drown in his own reflection.

The mania for domination and greed

The mania for domination is caused by a fear of being dominated by others and intend to use force over the others. Kant distinguishes this passion from the indirect art of domination, such as the female art to dominate through charm: "Men dominates by the use of force, women dominate by the use of charm" (Anth, AA 7: 274). In this claim Kant shows his sexist view of women who allegedly dominates by charm and beauty. This is not the mania for domination, that is the search for politically or social power. By having this power, men can dominate by the fear of others:

> It starts, however, from the fear of being dominated by others, and is then soon intent on placing the advantage of force over them; which is nevertheless a precarious and unjust means of using other human beings for one's own purposes: in part it is imprudent because it arouses opposition, and in part it is unjust because it is contrary to freedom under the law, to which everyone can lay claim (Anth, AA 7: 274).

Greed, the passion for possession or money can be explained by the fact that it is a power that people believe replaces the lack of every other power. Wealth can open all doors: "Money is the solution, and all doors thar are closed to the man of lesser wealth open to him whom Plutus favors" (Anth, AA 7:274).

Passions and the overcoming the propensity to indolence

Kant has a double standard regarding passions. On the one hand, passions are harmful, because they are considered cancerous sores for pure reason, on the other hand, in *Idea for a universal history with a cosmopolitan aim*, passions are responsible for the overcoming of indolence.

The fourth proposition of the idea reads as follows:

> The means nature employs in order to bring about the development of all its predispositions is their **antagonism** in society, insofar as the latter is in the end the cause of their lawful order (IaG, AA 8: 21).

This antagonism is the unsociable sociability: man has at the same time a propensity to enter into society, and a resistance to do that. He has a tendency to become socialized, in order to feel more like a human being, but at the same time he wants to isolate himself.

The human beings resist the others, because they are driven by their own desires to get what they want. But – here is the turn- this resistance to the others, "driven by ambition, tyranny and greed" is a path from crudity to culture. If men live in "perfect content, contentment, and mutual love", they will live like beasts and will not develop their rational nature.

> Thus happen the first true steps from crudity toward culture, which really consists in the social worth of the human being; thus all talents come bit by bit to be developed, taste is formed, and even, through progress in enlightenment, a beginning is made

> toward the foundation of a mode of thought which can with time transform the rude natural predisposition to make moral distinction into determinate practical principles and hence transform a pathologically compelled agreement to form a society finally into a moral whole (IaG, AA 8 : 20).

Kant praises here the passions, as a drive to develop their humanity:

> Thanks to the nature, therefore, **for the incompatibility, for the spiteful competitive vanity, for the insatiable to possess, or even to dominate!** For without them all the excellent natural predispositions in humanity would eternally slumber undeveloped. The human being wills concord; but nature knows better what is good for his species: it wills discord. He wills to live comfortably and contentedly, but nature wills that out of sloth and inactive contentment he should throw himself into labor and toils, so as, on the contrary, prudently to find out the means to pull himself again out of the latter (IaG, AA 8: 21, Fourth proposition).

In the *Idea for an Universal history with a cosmopolitan aim*, passions work like an incentive to overcome sloth, and to make men more active to work and develop their rational predispositions. While analyzing the fourth proposition of the *Idea*, however, a reader of the *Groundwork* and of the *Anthropology* has a little embarrassment related to this account of passions. In the *Groundwork*, Kant claims that inclinations are usually impediments for the moral actions. In the *Anthropology*, although both affects and passions are considered illnesses of the mind, passions are undeniably worse. How can these cancers of the mind become suddenly a good impulse — the only one — to go from crudity to culture?

Are passions evil or not? Are passions good for society or are they the source of irrationality?

I claim that we have two distinct answers in Kant. One is based on the above quotes of the Idea. According to this answer, passions are good for society, they are not the cause of their irrationality, but are the possibility of the development of the natural predispositions of human being.

The other answer comes from *Anthropology* and — more radically — from the *Religion*, where he claims that passions are the evil we should fight against. More than inclinations, or affects, the evil principle refers to passions "which wreak such great devastation in his originally good disposition" (RGV, AA 6:93). Passions are considered here the true evil. Stoics are wrong when they fight against inclinations. Evil cannot be sought in inclinations, but "only in that which determines the power of choice as free power of choice (in the first and inmost ground of the maxims which are in agreement with the inclinations)" (RGV, AA 6: 59).

In this text, he also claims that passions only arise when someone is among other human beings:

> If he searches for the causes and circumstances that draw him into this danger and keep him there, he can easily convince himself that they do not come his way from his own raw nature so far as he exists in isolation, but rather from the human beings to whom he stands in relation or association (RGV, AA 6: 93).

Passions are social and come from the comparison among human beings:

> He is poor (or considered himself so) only to the extend that he is anxious that other human beings will consider him poor and will despise him for it. Envy, addiction to power, avarice, and the malignant inclinations associated with these, assail his nature, which on its own is undemanding, as soon as he is among human beings. Nor it is necessary to assume that these are sunk into evil and are examples that lead him astray: it suffices that they are there, that they surround him, and that they are human beings, they will mutually corrupt each other's moral disposition and make one another evil (RGV, AA 6: 94).

Here Kant stresses the social nature of passions, since they are born out of comparison between people. It is not about evil people corrupting a good person, but the corruption of

the good person by herself through comparison and fear to be dominated or despised. Then, someone will become greedy, because he fears that he will be despised because he is poor or dominated because he does have any power.

Also, in the *Religion*, he claims that the civil society is not enough to heal evil, but we need an ethical society. The only way to have the victory of the good principle over the evil one is through "a society in accordance with, and for the sake of, the laws of virtue" (RGV, AA 6: 94).

This society is an ethical society, that differs from the civil society:

> An association of human beings merely under the laws of virtue, ruled by this idea, can be called an ethical and, so far as these laws are public, an ethico-civil (in contrast to a juridico-civil society), or an ethical community. It can exist in the midst of a political community and even be made up of all the members of the latter (indeed, without the foundation of a political community, it could never be brought into existence by human beings) (RGV, AA 6:94).

Only an ethical society can lead to full social rationality and the juridico-civil society is a necessary, but not a sufficient condition to that.

Conclusion

Is there a unique understanding of passions in Kant? Are they good or evil for progress? What is the true Kantian conception?

Perhaps we could understand the positive aspects of passions only as conjectural beginning of human history where passions had a good function as something that take men out of their natural inactivity. Once the civil society is stablished, passions are impediments to achieve an ethical society, since they usually work against the moral law. Vainglory, ambition and greed lead to irrationality in human social life. Only an ethical community can show the way to a full rational society.

Bibliographic References

The works are cited according to the Academy Edition:

Kant, Immanuel. *Gesammelte Schriften*. Ed.: Bd. 1-22 Preussiche Akademie der Wissenschaften, Bd. 23 Deutsche Akademie der Wissenschaften zu Berlin, ab Bd 24 Akademie der Wissenschaften zu Gottingen. Berlin: Walter de Gruyter & Co., 1900-.

KANT, I. *Anthropologie in pragmatischer Hinsicht. In: Gesammelte Schriften*. Bd. 07.

KANT, I. *Idee zu einer allgemeinen Geschichte in Weltburger Absicht. In: Gesammelte Schriften*. Bd. 08.

KANT, I. *Metaphysik der Sitten. In: Gesammelte Schriften*. Bd. 06.

KANT, I. *Die Religion innerhalb der Grenzen der blossen Vernunft. In: Gesammelte Schriften*. Bd. 06.

KANT, I. *Vorlesungen uber Anthropologie/Mongrovius. In: Gesammelte Schriften*. Bd. 25.

Recebido 11/2022 | Aprovado 12/2022

The progress of reason as an end of social rationality in Kant?

[O progresso da razão como um fim da racionalidade social em Kant?]

Sulamith Weber[1]

Universität Vechta (Vechta, Alemanha)

DOI: 10.5380/sk.v21i1.91539

Abstract

The essay investigates whether from a Kantian perspective, we can gain insights for a deeper understanding of the concept of social rationality. One strength of Kant's approach is to distinguish from a concept of mere formal rationality an emphatic account of reason as a faculty of human beings, anchored in human nature and destined to be fully developed. By distinguishing two dimensions in which the progress of reason is seen by Kant as the highest aim of nature – first, as ultimate end of nature, then, as final end of nature –, we can achieve an extended concept of social rationality, denoting not only a means for the promotion of the progress of reason but also the heuristic idea of a social reason itself, a kind of social reasonableness.

Keywords: social rationality; progress of reason; Kant; teleology; reasonableness.

1 sulamith.weber@posteo.de

At first sight, this title might cause irritation. Even before determining what exactly should be understood by the terms "progress of reason" and "social rationality" – if we roughly consider the first expression to designate some kind of process by which thinking beings acquire more and more rational skills and the second one equally roughly to refer to some state of society that we desire – doesn't it seem more logical to put it the other way round: that the progress of reason were the means by which social rationality is realized.

A different kind of objection could point to the apparently tautological character of any means-end relationship between something called "reason" and something called "rationality", considering that both terms derive from the Latin word *ratio*, standing for reason, calculation, information, relation, and principle.

With this essay, I try to respond to both objections. While the first one will be explicitly discussed, I hope to implicitly rebut the second one as well.

I will start with some reflections on the concept of social rationality (1). Then, discuss the idea of the progress of reason in some of Kant's works (2), before finally, suggesting an account of social rationality in a Kantian perspective wherein the question of how to assess or measure social rationality, as well as the problem of methodological egoism in social rationality, will be briefly addressed (3).

1.

During the 20th century, the concept of rationality came to supersede the concept of reason in large parts of philosophical discourse. To underline this, many examples could be given from the instrumental rationality at the center of Max Weber's sociology (cf. Weber, 1978, p. 24–26), John Rawls's "deliberative rationality" in his *A Theory of Justice* (cf. Rawls, 1999, p. 365–372), *as well as*, Jürgen Habermas' belief, in his *Theory of Communicative Action*, that all modern, post-metaphysical tendencies of philosophy converge into the striving for a theory of rationality (cf. Habermas, 1995, p. 16) – just to mention a few.

This notion of rationality can be said to correspond to the notion of reason similar to the way "morality" corresponds to "morals" or "modernity" to "modern age." That is to say, in all cases where newer expressions increasingly get implemented into our discourses, expressions that more or less recently superseded their respective older, traditionally more loaded words. It is noteworthy that these newer expressions tend to be built from the adjective (rational, moral, modern) deriving from the corresponding noun (reason, morals, modern age) that appeared in the older expression.

On the one hand, this conceptual shift takes account of changes that occurred in our self-conception as human beings, changes inspired both by natural and social sciences, and thereby guarantees that we can maintain the discourse about what it means for us to be reasonable or rational despite our becoming skeptical of whether there is a distinguishable human power that *makes* us reasonable or rational. This aspect is highlighted and explained by the German philosopher Herbert Schnädelbach, whose philosophical work has a main focus on the philosophy of rationality:

> A new term came to replace an old one. 'Reason' sounds antiquated, 'rationality' not at all. Talking about reason can be suspected of engaging into metaphysics, while the subject of rationality seems to be scientifically sound. 'Reason' is evocative of terms like consciousness, soul, spirit, i.e., of psychological entities that are of no use anymore for today's psychology. Rationality, however, is accepted by psychology as its object of inquiry even if psychology understands itself as pure behavioral science. Rationality, then, is the empirically measurable dispositional property of persons and systems. This is how rationality replaced reason (Schnädelbach, 1984, p. 8, my

translation).

From this point of view, we can say that the conceptual shift from reason to rationality updated the older concept for our modern purposes: By talking about "rationality", the concept of reason is freed from problematic metaphysical burdens, and thereby allowing us to continue philosophizing about it without needing to worry.

On the other hand, if we superficially reflect on what happens in the transition from the concept of reason to the concept of rationality, then we notice that something might get lost here: Max Weber's instrumental rationality, to take up this example again, is a purely formal rationality. In the chapter "Basic sociological terms" of his book *Economy and society*, the notion of reason itself is not once mentioned (cf. Weber 1978 3–62). John Rawls, who, with his "deliberative rationality," also considers a formal concept of rationality, shows himself as even more skeptical when he wishes to reduce the remaining concept of rationality to principles of rational choice instead (cf. Rawls 1999 361). Even Karl-Otto Apel, who provides one of the most rigorous and ambitious theories of rationality in philosophy today, has a concept of rationality that is, although universal, merely methodical and, therefore, in a sense, formal: While arguing against a methodological rationality, he favors methodical rationality (*methodische Rationalität*) instead (cf. Apel, 1973, p. 1).

We only need to turn to Kant to see that reason can very well be thought of in science and even in philosophy without necessarily 'substantializing' it into a specified metaphysical entity. In Kant, we find an emphatic account of reason as a faculty (*Vermögen*). The specifics of this Kantian approach manifest themselves in his paying attention to the fact that it is indeed problematic to want to study reason merely as some object of investigation among others – because reason is at the same time, and necessarily, the subject of all investigation. Thus, before we can ask what is the objective nature of reason, we must – if we follow Kant – make sure we know the subjective nature of reason, that is, its scope and its limits. Starting from this insight, Kant further develops his transcendental account of reason, with the central idea that reason is limited, or rather, that it must limit itself to the scope of possible experience in the theoretical sphere, or else to a form of knowledge that is merely hypothetical, although necessary, in the practical sphere.

From this rough description we can see that if we think of reason as a faculty, the term "rational" as its derived adjective very generally can be understood in at least three ways.

(I) First, as it is applied to the objectively measurable *results* of reason, in the respective spheres (theoretical or practical). "Rational" would feature here as a predicate applicable to thoughts or actions respectively.

(II) Second, as a name for the still objectively judgeable *exercise* of reason. In this sense, the term appears in adverbial form, such as when we would speak of someone proceeding "rationally".

(III) Third, as applied to *reason as a subject itself*. By calling something "rational", we would then allude to a potential reasonableness that is present in a given situation and could be actualized if certain circumstances were met.

This third aspect is what I think is usually getting lost in our discourses on rationality. In focusing on the state and the procedure of rationality, which is not wrong, they do not yield the whole picture – at least if we agree that we can relinquish the idea of a substantialized reason without having to stop talking about reason altogether.

Having considered the concept of rationality in some detail, the question now arises what is meant by the term "*social* rationality". It seems to me that social rationality usually designates either a state of achieved social rationality (1) or a procedure of social rationalization by which

a state of social rationality is realized (2). Furthermore, it can be conceived of as either applying to a collective (a) or to individuals (b).

(1) Social rationality as a *state* can mean

(a) structures present in a well-working *collective* (such structures would be called socially rational insofar as they are well-grounded and correspond to or promote some ideal of our collective social life),

(b) the mindset of an i*ndividual person* who can harmoniously interact with others (such a mindset would be called socially rational insofar as it enables the individual to function as a member in a well-working society).

(2) Social rationality as a *procedure* can point to

(a) a process of coordination of *individuals* to make them form a prosperous, socially rational collective,

b) the education of the individual enabling her to increasingly act in accordance with others, that is to say, making her more and more socially rational.

It can be easily seen that social rationality as a state (1) corresponds to the first (I) of the earlier mentioned components of the term "rationality", and social rationality (or rationalization) as a procedure (2) corresponds to the second (II) aspect. While in rationality in general, we can at least imagine the third (III) aspect of the term to be present, this becomes doubtful when it comes to *social* rationality: Trying to think of a faculty of social reason seems to lead to inescapable ambiguities concerning the question whether we are talking about a collective or an individual faculty of reason; or if both, in what relation they stand.

I will come back to this problem later when I ask if we can, from a Kantian perspective, add to the two aspects of social rationality a third component, integrating a stronger relationship with the concept of reason into the concept of rationality. Trying to do this will bring along the idea that social rationality can (or, in the Kantian perspective, even has to) be thought of as having the purpose of what we can call the progress or development of reason, which is going to be the focus of the following section.

2.

We can find the idea of the progress of reason as a purpose in Kant in the essay "An answer to the question: What is enlightenment?" (1784). Here, Kant is taking up the main philosophical concern of his time, which is the problem of progress, and, more precisely, the question whether there is a progress of reason which could rightfully be called Enlightenment.

In order to understand the quite revolutionary character of the position Kant is assuming towards this problem of progress, we can take a brief glance at the philosophy of Enlightenment of his time. In most parts – with the famous exception of Jean-Jacques Rousseau – Enlightenment philosophy rests upon a very optimistic stance in questions of progress.

The idea of humanity striving continuously towards its own moral perfection as knowledge increases has been perhaps most famously developed by Nicolas de Condorcet in *Sketch for a Historical Picture of the Progress of the Human Mind* (1795 posthumous). But also the authors of the *Encyclopedia*, the Enlightenment project *par excellence*, were guided by this idea, as confirmed by Denis Diderot's announcement in 1765 of the publication of the final volumes of the *Encyclopedia* – a text where Diderot speaks unambiguously of the happy revolution that Enlightenment thinkers bring to human history. In both cases, in Condorcet and in the kind

of self-assuring procedure that the Enlightenment accomplishes with the *Encyclopedia*, the idea of progress is brought to and expressed in its highest form: as it is made entirely dependent on the ability and good will of rational beings to emancipate themselves. The objective progress in history follows naturally from the subjective progress of the human mind.

Kant, in "An Answer to the Question: What is Enlightenment?," certainly is not denying the idea of the progress of reason that is called Enlightenment. On the contrary, he takes his time to explain and to promote the Enlightenment agenda. But at the same time, he specifically does so in a new way: by taking one step back and asking what are *the conditions of possibility* of the progress of reason. In the essay, Kant's complex answer to this question is already foreshadowed by his alluding to human nature "whose original vocation lies precisely in such progress" (AA 8:39).

This idea of the progress of reason, or Enlightenment, as the vocation or the destination of human nature, we find in more elaborated form in the essay "Idea for a universal history with a cosmopolitan aim," which dates from the same year as the Enlightenment essay. In this text, Kant asks more explicitly how it is possible for us to conceive of a universal history in which we are embedded. Only in relation to such a conceivable universal history could a judgment be spoken concerning the question whether we live in a time of development to the better or to the worse.

It is not Kant's intention to present a certain version of history, either optimistic or pessimistic. Although we can say that he is finally taking sides in this question (opting for a concept of history as an advancement rather than a decline, in contrast to, for example, Moses Mendelssohn), interestingly, his starting point in the "Idea…" is exactly the opposite of an optimistic premise: The text starts with the observation that the ways humans act are far from being rational. From this observation, it is hard to see how it is that we indeed do have a history – it should seem much more probable that our individual actions, that are often thoughtless, therefore uncoordinated and often even opposing and negating each other, should result in a chaos.

Consequently, Kant asks how it is possible that we nevertheless conceive of a discernible history of mankind. His answer is, in short, that this is possible if and only if we think of nature as having the progress of reason for its purpose. Kant's argument has a transcendental and a teleological part, and I will explain one after the other, in order to make clear why Kant considers the progress or development of reason as an end of nature.

I begin with the transcendental argument, which has the classical structure of a transcendental argument in Kant: It starts with naming a problem or question – in this case, the concept of history –, analyzes the state we are in – consisting of a bundle of individual actions that to a large extent are not rational at all –, and it concludes by stating that in order to solve the problem in the given state we have to think about things a certain way, that is, we have to consider the state that we are in under a certain idea. In our case, this means: If there is to be something like history instead of a mere chaos of unrelated events, we must think of our individual actions as guided by a purpose, a purpose pursued by nature that concerns and involves mankind as a whole.

In the introduction to the first version of the *Critique of Judgment* (1790), we find this argument in more detail: First, Kant makes clear why a concept of history is of such a great importance to us – for otherwise, we could not trust our experience. Although the introduction to the first version of the *Critique of Judgment* focuses on the consequences of this problem in the theoretical sphere, we can easily extend the argumentation to the practical sphere: If we could not trust our actions to have an discernible effect in the sensible world and to be of a certain permanence, that is, to not in general be thrown over at the next opportunity by our fellow human beings, we would not have the motivation to act at all. But since we do act, we obviously

rely on the belief that effective acting is possible. As soon as we want to reassure ourselves about this belief, we must assume nature as acting purposefully. The two most important corollaries of this assumption are that 1. the laws that structure our experiences stay the same over time (in the theoretical sphere), and that 2. nature has a defined purpose with mankind, a purpose that forms the condition of possibility for our individual actions (in the practical sphere).

This idea of nature as a rational subject, as having purposes to itself, has the status of a heuristic yet necessary assumption. It is not certain in the way an analytic truth is certain, but it is necessary with the same necessity that is present in our having to consider our actions as potentially effective in order to act at all.

But until now, we only have the idea of *nature* as pursuing ends in general, and the specification that nature must have a purpose with mankind. What we do not have at this point of the argumentation is an answer to the question why *reason*, or the progress of reason, should be the highest purpose of nature.

This is being solved by another kind of argument that complements the transcendental one – an argument of a teleological kind. As it has already been established by the transcendental argument, regularities in nature (may they be physical events or the results of free actions) are to be considered under the idea of them serving a purpose of nature. We therefore have the heuristic idea of "nature [that] does nothing superfluous and is not wasteful in the use of means to its ends" (*Universal History*, AA 8:19). One of the regularities that we can find in nature is that human beings can be rational. Thinking of nature as indicated in the quotation just cited, we can take human nature as intended to be developed and fully realized.

In other words, the teleological character of the argument lies precisely here: that we observe human being's disposition of being rational and infer, by use of the heuristic idea of nature acting purposefully, from the presence of this disposition that it must serve a purpose. This purpose, in the case of a disposition, is easy to determine: it is realizing and developing the dispositional potential more and more towards its full realization.

In the "Idea…", the concept of end or aim (*Zweck*) remains rather vague: Kant calls the "development of all the predispositions in humanity" the "highest aim of nature" (AA 8:22).

This concept of purpose or end, which is key to understanding Kant's teleological argument, is deepened and explained further in the last section of the *Critique of Judgment*, the "Methodology of teleological judgment".

In §§ 83 and 84, after Kant has introduced the category of purposefulness as a heuristic assumption on the nature of nature, the concept of the "highest aim of nature" is given two different interpretations: in § 83, as an "ultimate end [*letzter Zweck*] of nature" and, in § 84, as a "final end [*Endzweck*] of nature". What is the difference between the 'progress of reason' being the *ultimate* end and the *final* end of nature?

§ 83 of the *Critique of Judgment* has the title "The ultimate end of nature as a teleological system", and Kant is very straightforward in the determination of this ultimate end of nature; he calls it culture: "[O]nly culture can be the ultimate end that one has cause to ascribe to nature in regard to the human species" (AA 5:431).

It is important to be clear on how the genitive in "ultimate end of nature" is to be understood here: as a *genitivus subiectivus* meaning that nature as a subject pursues culture as its ultimate end.

Kant defines culture as "[t]he production of the aptitude of a rational being for any ends in general (thus those of his freedom)" (ibid.), meaning, as we see, all kinds of ends, not necessarily moral ends. He further divides the concept of culture into two components that have to come together to bring about culture in its fullest sense, a culture of "skill" (ibid.) and

a culture of "discipline" (AA 5:432). I will come back to the Kantian notion of culture later. For now, it is only important to retain that the first way in which we can say that the progress of reason is an end of nature is by understanding the progress of reason as culture, as the development of the aptitude of human beings to be rational, that is, to set themselves purposes of all kinds.

Yet this is not the *full* development of the predisposition of reason. "Progress of reason", in the sense that is to be identified with the *ultimate* end of nature, is a *genitivus obiectivus* that could be reformulated as "progress *in* reason", meaning a mere immanent progress that does not allow for any criteria to judge the effects of this process itself. At this point, nothing guarantees that having a culture does mankind any good. Kant expresses this uncertainty when he states that "the end of nature itself, even if it is not *our* end, is hereby attained" (AA 5:432, my emphasis).

In § 84 of the *Critique of Judgment*, Kant turns to the final end or purpose of nature and explains what distinguishes it from any other end, including ultimate ends: "A final end is that end which needs no other as the condition of its possibility" (AA 5:434).

The ultimate end of nature was still an end *pursued by* nature. There is no reason why such an end, even if it is the "ultimate", most general one, should not also feature as means in other contexts. As they are both still of the natural sphere, we can imagine culture and the human race being used to bring about ends different than themselves.

If we are talking about the "final end of nature", this is now a *genitivus obiectivus*, meaning the purpose nature *itself* is serving, instead of the last element of the list of higher and higher ends pursued *by* nature. This can no longer be a natural purpose; otherwise we would again have an end that could at the same time be means for some other end. We are therefore necessarily looking for something that reaches *beyond* nature but, at the same time, it must be of the kind that it can be *promoted by* nature. There is, says Kant, only one thing on earth that inhabits both spheres, the natural one and the sphere of ends, or of free action, and this is man as a moral being: "[O]nly in the human being, although in him only as a subject of morality, is unconditional legislation with regard to ends to be found, which therefore makes him alone capable of being a final end, to which the whole of nature is teleologically subordinated" (AA 5:435–436).

The "progress of reason" can now be understood as a *genitivus subiectivus* as well: Reason is not only blindly developing further in an immanent process of unfolding its potentials that have been preinstalled by nature, as we have seen above, but it is at the same time the subject of this process, which means that it has to orient and determine itself while developing. In an ideal sense, we can now speak, reading the genitive subjectively, of the progress of reason as *our* end. "Ideal" because it is of course in no sense guaranteed by nature that human beings should act from their ability of being moral all the time.

Thus, for Kant, that towards which we as moral subjects orient ourselves cannot be what seems desirable to us – for this is arbitrary, and would lead, as pointed out in the "Idea…", to a chaos of individual actions inhibiting and fighting each other. Instead, as we learn from the *Critique of Practical Reason* (1788), that towards which we orient ourselves in moral acting must be *the condition* of everything that could be desirable on Earth, and this is virtue, the ultimate good (*oberstes Gut*).

3.

It has become clear that the progress of reason can be taken as end or purpose of nature

in two different ways – I summarize what I have said in the second section:

The full development of the *predisposition of reason* in mankind is *the ultimate end of nature* (where nature is, in a heuristic assumption, the subject that sets purposes to itself and, mankind, as the carrier of reason, is the object or medium of the process). At the same time, this is only the precondition for the full development of *reason itself* that can take place in a second step: the progress of humanity itself towards morality as *the final end of nature* (nature featuring in this part of the process as means; while humanity or reason itself is the subject of the development).

What has not yet become clear is what role social rationality could play in the context of the progress of reason as an (ultimate or final) end of nature as described above. In the title, I suggested that social rationality, from a Kantian perspective, has the progress of reason for its purpose. Now, in analogy to what we have seen earlier, we could understand this in two ways: either with the progress of reason (objectively read genitive) as the *ultimate end* of social rationality (subjectively read genitive), in the sense that social rationality as the 'subject' of the process promotes the development of the predisposition of reason in humanity, or with the progress of reason (subjectively read genitive) as the *final end* of social rationality (objectively read genitive), in the sense that social rationality itself serves as an end that is beyond itself, i.e., reason as the true subject of morality. In the following, I want to shortly examine both ways of, as I claim, a Kantian manner to think of social rationality, starting with the second one, the progress of reason as the final end of social rationality.

As developed earlier, social rationality from a conceptual point of view could be understood in at least three ways. It seems to me that the first two aspects of the term, social rationality as a state and as a procedure to bring about such a state, which are explicit and well-represented in our actual discourses about social rationality, can both be found in the Kantian concept of culture.

We have seen that Kant defines culture as "[t]he production of the aptitude of a rational being for any ends in general" (KU, AA 5:431). The German term for "production", *Hervorbringung*, seems deliberately chosen by Kant as to cover both the procedure and its result. Kant could have written "das Hervorbringen", which, like "producing" in English, could have served to designate only the procedure or, if he had wished to only designate the result, he could have done without a term like *hervorbringen*, simply defining culture as the aptitude of having purposes. For reasons of clarity, from now on I am going to write "cultivation" if I want to point to the procedural character of culture, "cultivated state" when it is about the result of cultivation, and "culture" if I intend to designate both. Culture for Kant is understood as a cultivation of *technical* predispositions aiming to develop what Kant names "skill" (*Geschicklichkeit*) and a cultivation of *practical-moral* dispositions leading to "discipline" (*Zucht*). The corresponding cultivated state, Kant determines as to include art and science.

As I already indicated, the state brought about by cultivation is by no means without conflict nor is it an especially moral one. Kant, in the cited chapter of the *Critique of Judgment*, says that art and the sciences do not make human beings more moral, but only more civilized (cf. AA 5:433). However, this begs the question of why we should have social rationality at all if it is in such a high degree permeated by social irrationalities of all kinds. The Kantian answer is the one we have seen for the case of nature in general, we only need to apply it to social rationality as a state and as a procedure, both belonging to the natural sphere: Mediated by the heuristic but necessary idea of nature as a purposeful subject, social rationality in these first two aspects of the concept has for its final purpose the progress of reason itself in its orientation towards morality.

What has not yet been discussed is the question of the ultimate end or ends of social rationality and, in connection with this, the question of how social rationality in a society or an individual could be assessed or measured. Of course, it is no longer a question here of social

rationality as a state or procedure, but in the possible third aspect of the concept: as a potential of a certain kind of reason itself, i.e., a kind of social reason.

As long as it is predefined what social rationality is – a *certain* formal procedure, a *certain* desirable state – it is of course easy, if not trivial, to measure. The real question of measurement arises precisely from the questioning of such fixed predefinitions of what should or should not count as social rationality, asking ourselves: How can we say if and to what degree a given situation is socially reasonable or, equivalently, how can we measure social rationality taken in its third aspect?

As a matter of fact, it is merely a heuristic assumption to talk of social rationality as potential social reasonableness, an evolving subject of social reason. Still, we have to make this heuristic assumption as soon as we want to assess the degree of social rationality realized in a given situation.

Again, we can apply what Kant has said in his texts explicitly for the case of nature to social rationality as a part of nature, more precisely, as part of "those [natural] predispositions [of man] whose goal is the use of reason" (*Universal History*, AA 8:18). We can then look at and judge our social reality under the heuristic idea of a social reason taken as a subject which has ends, and, above all, the ultimate end of realizing itself fully in our social reality.

Conversely, this means that if we assess the question of measurement of social rationality, then there still must be *empirical* criteria by which we can judge our social reality, even under the idea of a social reasonableness that realizes itself in the world. This is important to retain, because here we can see that a Kantian theory of social rationality – at least in light of my reconstruction – would not be purely formal, which contradicts a very widespread criticism of Kant.

Finally, I would like to briefly comment on the question I raised in the first section, whether such a social reason would be individual or collective, and add some further thoughts on the problem of methodological egoism in social rationality. I think we could now say that social reason would have to be both individually and collectively instantiated. As *potential* social reasonableness, it is a "natural predisposition" of each individual and has to be realized as such, but it is only in the species, not in the individual, that it shall develop itself *completely*, as Kant writes in the second paragraph of the "Idea..." (AA 8:18).

From here, it becomes clear that Kant's account of social rationality is not one that allows for any interpretation that bases social rationality on individualistic and egoistic motives alone, aligned and harmonized with each other by some external influence – be it some postulated effective force (one could think of Adam Smith's thesis of the "invisible hand") or a merely formal procedure, as in some of the accounts of rationality mentioned earlier in this essay.

In order to measure or increase social rationality in our societies, we cannot solely rely on dynamics belonging to the development of what Kant calls cultivation, whereby individuals are still incapable of setting themselves purposes in general. These purposes have to become moral ones for the full development reason as a natural predisposition in human beings, and, accordingly, for a higher social rationality in all three aspects of the term. The judge or touchstone of this development towards morality, as we have seen, can only be reason itself as it is developing – there is no external or internal authority to determine what the outcome of this process has to look like. In the meantime, and from an anthropological perspective, Kant advocates for "the opposite of egoism [that] can only be pluralism" (*Anthropology*, AA 7:130). Kantian pluralism is about regarding oneself not as isolated and exceptional (as the logical, the aesthetic and the moral egoists do, each in their respective area), but as a member of the species in which reason is to be developed or as "a mere citizen of the world" (ibid.).

Bibliographic References

APEL, K.-O. *Transformation der Philosophie*, Frankfurt: Suhrkamp, 1973.

HABERMAS, J. *Theorie des kommunikativen Handelns*, vol. 1, 4th edition, Frankfurt: Suhrkamp, 1995.

KANT, I. *Practical Philosophy* (The Cambridge Edition of the Works of Immanuel Kant), ed. by M. Gregor. Cambridge: Cambridge University Press, 1996.

KANT, I. *Critique of the Power of Judgment* (The Cambridge Edition of the Works of Immanuel Kant) ed. by P. Guyer, transl. by E. Matthews. Cambridge: Cambridge University Press, 2000.

KANT, I. *Anthropology, History, and Education* (The Cambridge Edition of the Works of Immanuel Kant), ed. by R. Louden & G. Zöller. Cambridge: Cambridge University Press, 2007.

RAWLS, J. *A Theory of Justice*, revised edition. Cambridge: Harvard University Press, 1999.

SCHNÄDELBACH, H. *Rationalität*, Frankfurt: Suhrkamp, 1984.

WEBER, M. *Economy and Society*, ed. by G. Roth & C. Wittich. Berkeley/Los Angeles/London: University of California Press, 1978.

Recebido 11/2022 | Aprovado 02/2023

The Kantian view of dark representations and their function in practical life, according to the anthropological notes of the Critical Period

[A visão kantiana das representações obscuras e sua função na vida prática, de acordo com as notas antropológicas do Período Crítico]

Luciana Martínez[1]

Universidade Federal do Paraná (Curitiba, Brasil)/Universität Vechta (Vechta, Alemanha)

DOI: 10.5380/sk.v21i1.91540

Abstract

This contribution examines the doctrine of dark representations as it is presented in the anthropological annotations of the critical period. The explanation of such representations is analysed, taking into account the analogies used for this purpose. Three kinds of manifestations of such representations in everyday life are then considered: our connection with death and fear, the development of certain feelings, such as respect, and their effects, and the way we produce philosophical knowledge.
Keywords: Dark representations; Fear; Feelings; Honour; Philosophy.

Resumo

Esta contribuição examina a doutrina das representações obscuras tal como ela é apresentada nas anotações antropológicas do período crítico. A explicação de tais representações é analisada, levando em conta as analogias usadas para esse fim. Três tipos de manifestações dessas representações na vida cotidiana são então considerados: nossa conexão com a morte e o medo, o desenvolvimento de certos sentimentos, como o respeito, e seus efeitos, e o modo como produzimos conhecimento filosófico.
Palavras-chave: Representações obscuras; Medo; Sentimentos; Honra; Filosofia.

1 luciana.mtnz@gmail.com. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, within the scope of the project *Kantian Perspectives on social irrationality*.

Introduction

What does it mean that our representations are dark?[2] And are there such representations in our minds? These questions have been of interest to many modern philosophers, and Immanuel Kant has been no exception. All along the Kantian sources we can find several references to this issue. This does not imply that the philosopher's position has been univocal throughout his life. References to dark representations in sources from different disciplines are not unambiguous either. Logical texts differ from critical publications and the anthropological corpus in their approach to such representations. The pre-critical anthropological corpus has already been studied in a previous article and the logical corpus is examined in another essay. In this paper, we will focus on the anthropological sources of the critical period, i.e. the 1780s.

For his Anthropology lectures, Kant used the section of Alexander G. Baumgarten's *Metaphysica* devoted to empirical psychology. Baumgarten points out that a person thinks sometimes distinctively, sometimes confusedly. The characteristic feature of confusing thinking is that in it we do not distinguish the notes of what we represent. When we do distinguish the notes of our representation, then we think with distinction. That which we cannot think with distinction remains instead a dark representation. Obscurity is a feature of our sensible representations, whose notes we have not distinguished through thought. Our thinking is able to give distinction to our representations, as long as it distinguishes their notes. Baumgarten states that there are dark representations in the soul. In the text, he explicitly links the obscurity of certain representations with the position of one's body in the universe. Therefore, the soul is conceived by him as a force (*vis*) that represents the universe according to the position of the corresponding body in it.[3]

In his lectures, Kant developed an explanation of the subject that differs significantly from Baumgarten's position. Besides, over the years his thoughts seem to become more precise and orderly. On the one hand, he establishes a specification of concepts. To clarity (*Klarheit*) he opposes obscurity (*Dunkelheit*); to distinction, indistinction (*Deutlichkeit*, *Undeutlichkeit*). Moreover, he abandons the identification of confused (*verworren*) knowledge with sensible knowledge and the explanation of all of them on the basis of the position of the knowing body in the universe.[4] Throughout all the anthropology notes currently available, however, we notice that he maintains the thesis that *there are* dark representations, which are unconscious representations.

As stated, then, in this paper we will examine some aspects of dark representations, in accordance with the considerations contained in the anthropology lectures of the critical period.[5] The texts taken into consideration are, in particular, the anthropological notes *Menschenkunde* from the winter semester 1781-1782, *Mongrovius* from 1784-1785 and *Busolt* from 1788-1789. All these notes are published in volume 25 of the Academic Edition of Kant's works.[6] The Kantian doctrine of dark representations in the pre-critical period has been the topic of a previous research (Martínez, 2014). In the present paper we develop the conclusions of a subsequent research, focusing on the critical period of Kantian philosophy. As in the pre-critical period, these annotations reveal the philosopher's commitment to the existence of dark representations. Moreover, Kant goes on to argue that most of our mental life involves this kind

2 In this article, we will mainly use the term "dark" to translate the German term "*dunkel*". The term "*obscure*" could equally well be used, but I think it should be reserved for translating the passages in which Kant expresses "*obscure*", which by all means seems to be interchangeable with "*dunkel*".

3 Baumgarten, *Metaphysica*, §§505-513, en AA 15:5f.

4 Cf. Sánchez Rodríguez (2012, 2013).

5 This article focuses on pragmatic-psychological explanations of unconscious representations and their manifestations in practical everyday life. The investigation of the treatment of the relation of our unconscious representations to Kant's practical critical philosophy has been developed by P. Giordanetti (2012).

6 In this paper, Kant's works are quoted from the academic edition, according to the standard conventions explained in the journal *Kant Studien*.

of representation.

Similarly, in the later lectures we find some of the images and metaphors that were already present in the early classes. Through these rhetoric resources, the nature of dark representations and their link with the rest of our mental life is explained. We refer in particular, on the one hand, to the metaphor of our mind as a map that has only some parts enlightened, and on the other hand, to the explanation of the difference between clear and dark representations by means of some appreciations on the observation of the Milky Way.

In the lecture notes of the critical period we find, moreover, some novelties with respect to the previous period. First, Kant distinguishes two ways in which we are linked to dark representations. On the one hand, they constitute us. We are a set of dark representations, which condition the way we live. On the other hand, Kant points out that we play with dark representations. Additionally, in the later lectures the philosopher goes further into the examination of the visible effects of our dark representations on actual life and decisions. In the notes of the pre-critical period we have found a rather detailed study of the very notion of dark representations, criticisms of Baumgarten, explicit references to Hume and Locke, and, especially, a linking of dark representations with our feelings of pleasure/displeasure, with the doctrine of prejudice and with their function in knowledge. In the notes from the critical period, on the other hand, the approach to the subject is briefer and the references are more related to practical, everyday aspects of our lives, as we shall see.

In the following, we will first analyse the meaning of dark representations and the Kantian arguments concerning their existence and the way in which they are reached. Secondly, we will explain in what sense "we are a play of dark representations" and in what sense Kant considers that "we play with dark representations".[7] Finally, we will deal with three aspects of our practical daily life in which it is shown that dark representations constitute us. Hence, after reviewing the Kantian presentation of dark representations, we will examine the following effects of them: our link with the issue of death, the relation between lies and honour, and the philosophical task as a procedure of analysing representations.

1. The notion of dark representations and the claim that they exist.

According to student notes from a variety of disciplines, taken in the critical period, dark representations are those of which I am not *directly* conscious, but of which I may be *indirectly* conscious through inference.[8] They present themselves in mere sensation and we become conscious of them through inferences.[9] But what does it mean to be conscious (*bewußt*) of a representation? Unlike in the pre-critical lectures, the explanations in the notes of the period under review are brief.[10] Within the framework of the pragmatic-anthropological approach, the notion of self-consciousness interests Kant for the purpose of explaining the problems involved in focusing attention (*Aufmerksamkeit*) on oneself, rather than on other objects. One of these problems, for example, is hypochondria.[11] Kant explains that we can distinguish between subjective and objective consciousness. Our consciousness is subjective when we direct our

7 V-Anth/ Mensch, AA 25: 869; V-Anth/Mrong, AA 25: 1221; V-Anth/Bus, AA 25: 1440.

8 V-Met/Mrong, AA 29:879. V-Log./ Wien., AA 24: 805; V.- Log./ Pöl., AA 24: 510.

9 V-Anth/Mensch, AA 25: 867.

10 In this paper we will not deal with the critical notion of consciousness, neither in its theoretical use nor in the practical context. For a general explanation and a summary of some comments, see Kitcher, P., in Willaschek et al. 2010, 281-285. In the framework of logic courses, consciousness is presented as a "representation of our representations" (V-Log/Dohna-Wund, AA 24:701), or as the representation that we have representations. The consciousness of representations is, however, a starting point for logic, which does not deal with dark representations. Cf. V-Log/Wien, AA 24: 805.

11 V-Anth/ Mensch, AA 25: 862.

thoughts to our existence and to our mind itself; it is objective when we direct it to other objects.[12] According to the notes, he considered it healthier and more profitable to direct our attention to other objects.[13]

After exhibiting the disadvantages of concentrating attention on oneself, the professor dealt with attention to others. He moved from subjective to objective consciousness of objects. It is pointed out that we have both clear and dark representations of things, and that Kant was interested in concentrating on the latter.[14] In the notes we can read that dark representations are unconscious (*unbewußt*) and that we do not know (*wissen*) that we have them.[15] However, Kant would have considered it to be a mistake to assume that those representations of which we are not conscious are nothing to us. This is why he opposed John Locke and argued that most of our mental life is made up of dark representations. According to Kant, moreover, these representations are the very foundation of the others. To have representations and yet not to be conscious of them seems a contradiction; for how could we know that we have them if we are not conscious of them? Kant considered this to be the position of Locke,[16] who therefore would have rejected the existence of such representations. But as stated, Kant would have argued that we can be *indirectly* conscious of having a representation, even if we are not directly conscious of it.[17]

Now then, how is the difference between being conscious of a representation immediately and through reasoning to be understood? Kant suggests an explanation through certain images. For him, the dark representations are in us in the same way that the stars of the Milky Way were visible to the eyes of the ancients when they saw it without the support of instruments. If we look at the night sky, without using any equipment, we can see a blurred band of white light. In fact, its name comes from the mythology of the Greek tradition, for which the origin of this light was the milk spilled by its mother when she breastfed Hercules. To anyone looking at the sky without the aid of a telescope, the Milky Way looks like some white light. Only with the development of the telescope, Galileo Galilei was able to see that such a band of light actually consisted of a cluster of stars. The human eye does not identify the stars that compose it and they are therefore not perceived. However, their effect is presented to the eye. What the ancients saw was precisely the effect of something that was certainly there without being noticed. Our link to our dark representations is analogous to the link the ancients had with the stars of the Milky Way. Our dark representations are in our minds and we have access to their effects. But we are not conscious of them themselves. When we represent ourselves and our representations, we do not consider our dark representations. However, their effects are part of our mental life.[18]

On this point, before advancing our argument, I would like to raise two remarks. First of all, if, as we pointed out, Kant considers that dark representations become accessible through their effects, the examples he provides in this respect undergo some changes in the critical period, according to the available testimonies. One of the most noticeable modifications is the omission of prejudices. Instead of referring to them, in the lectures of the 1980s Kant seems to have chosen to mention the relation of dark representations to certain feelings that affect practical life. In addition, the treatment of analytic philosophy and its relation to the ideas of Socrates is developed in greater detail. As mentioned above, these issues will be addressed

12 V-Anth/Mrong, AA 25:1219.

13 V-Anth/Mrong, AA 25:1219.

14 V-Anth/Mrong, AA 25:1221.

15 V-Anth/Mensch, AA 25: 867.

16 Strictly speaking, in the *Essay* (II.xxix.5), Locke develops a slightly different line of argument, which is more in line with objections that Kant presents in other texts, such as KrV B415n. Locke says that the idea we have in our mind is the representation we have and it differs from the others in so far as it is that way. That is, if we have two ideas which we cannot distinguish from each other, then we have no criterion for distinguishing them and it makes no sense to hold that they are two different ideas.

17 Anth, AA 07: 135. V-Met/Mrong, AA 29: 879.

18 V-Anth/Mrong, AA 25: 1221.

in the following sections. The second point I would like to stress is that in the available notes a semantics of causes and effects is played out to explain the relationship between dark representations, which are described as causes or foundations (*Gründe, Quelle*) and conscious life, which is presented as an effect of them.[19]

Thus, for Kant, that there are dark representations is something we know indirectly, insofar as we can appreciate their effects. But the philosopher considers, moreover, that these dark representations constitute the greatest part of our mental life. Our mind is described in the notes as a map in which only certain parts are enlightened. Most of its surface remains obscured. As can be seen from this analogy, dark representations are in a majority and are also fundamental to our mental life. Besides, the image of the map that represents our mind is a constant that we find already in the notes of the first lectures of our philosopher, in the 1770s.

2. The dual *Spiel* of the dark representations

Our relation to our dark representations is, for Kant, twofold. This point is a novelty of the annotations of the critical decade. The differentiation that we will detail here is not developed in the pre-critical anthropology lectures, but is present in all the notes from the 1780s that are currently available. On the one hand, as we have stated, numerous dark representations constitute our mental life. This is reflected in certain attitudes. Thus, for example, some of our fears and beliefs, as well as some of our feelings, are based on dark representations. We will deal with this feature, which Kant enunciates by saying that "we are a play of dark representations" in the next section of this work.

On the other hand, Kant points out that in addition to "being a play of dark representations", we "play with dark representations". This aspect is not the main focus of our research, so we will comment on it in a general way. In this case, Kant refers to certain modes of communication in which we deliberately do not make explicit everything we want to say. Kant mentions that, in particular, for example, we are deliberately obscure when referring to aspects of life in which we resemble animals.[20] One example is the talk about sex. This discourse is full of metaphors and paraphrases that obscure what is referred to. There is a kind of play between what is said and what is not said, sometimes for reasons of discretion and often for reasons of satisfaction.[21] The suggestiveness and metaphors are pleasing to us. That is why we use literary tropes and pictures.

The anthropological treatment of dark representations thus has two dimensions. On the one hand, a myriad of dark representations are in our minds and affect, without our being conscious of it, the way we think and act. On the other hand, these representations fulfil significant social functions insofar as we make use of them for playful purposes. We deliberately take advantage of the difference between what is said, what is conscious, and what is suggested and unspoken. In this case, there seems to be a dark side to communication and interaction. This aspect does not have to do with the subject's lack of consciousness, but with deliberately omitted presuppositions and premises of his speech. As anticipated, in this paper we will concentrate on the first aspect of dark representations. Then, in particular, we will examine some effects of such representations on the action of the subjects.

19 V-Anth/Mensch, AA 25: 869.

20 V-Anth/Dohna, Ko 80.

21 V-Anth/Mensch, AA 25:872-874; V-Anth/Mrong, AA 25:1224.

3. Dark representations and fear

A series of examples illustrating the relevance of dark representations in our everyday life is recurrently mentioned in the anthropological annotations of the critical period. At this point, a topic from the philosophical tradition is taken up, which is that of our fear and preoccupation with death. Already in his *letter to Meneceus*, Epicurus had pointed out that death, as a deprivation of feeling, is nothing to us and that for this reason it is meaningless to worry about it. In death, there is simply no pain. The approach to the subject in Kant's lectures, however, is from a different perspective, namely that of empathic concern for the other who has died.

Unreasonable though it is, it is argued, we regret when someone young dies. Reason suggests to us that the dead person feels nothing. However, dark representations lead us to imagine him in the grave and feel pity. Our conscious, rational thinking teaches us that the person who has died does not experience any discomfort, moreover, he does not experience anything. But something unconscious in us motivates enormous grief and expresses itself in feelings of sorrow. These feelings are the conscious manifestation of causes that sink into the darkness of the mind.[22] Similarly, we worry about the stay of the dead body, as if it were relevant in any respect. For that body, however, everything is totally indifferent, even the place it occupies. The concern of those who reflect on where to deposit the corpse is not based on reasoning and rather seems to go against the indications of reason.[23]

Kant says something similar about the origin of some of our fears. When we climb a very high and well-built tower, we have an irrational fear of falling.We are conscious that statistics make our downfall unlikely and that construction is designed to provide us with security. However, such reasoning is not enough to overcome the fear, which presses upon us. Our fears are not based on reasons. Instead, they are grounded in unconscious representations, of unknown origin, which we cannot control.[24] Kant also mentioned other fears whose source seems to have been dark representations. For example, if we see a weak bridge over a river and see other people crossing it successfully, it may well be that we are afraid. Again, in this case, reason indicates that with caution we will come out safely, and the statistics point in the same direction. But our fear presents itself as inevitable.[25]

Thus, the line of argument suggested in the anthropology notes adopts a number of feelings as a starting point. Such feelings, like grief and fear, are manifested in our lives and suggest irrational motivations. If we examine their origin, we notice that they point to unconscious representations, which we only notice when we pay attention to them.

4. Dark representations, lies and honour

Another situation in which Kant warned of the incidence of dark representations is that of lying. Lying is, for our philosopher, irrational behaviour.[26] There are no motivations originating in our reason, which rather incites us to always remain truthful, in order to lie. The notorious Jorge Luis Borges has pointed this out with irony, according to the records of his colleague and friend Adolfo Bioy Casares:

22 V-Anth/Mensch, AA 25: 870.

23 V-Anth/Mrong, AA 25: 1223.

24 V-Anth/Mensch, AA 25: 870.

25 V-Anth/Mrong, AA 25: 1222.

26 It has already been noted that Kant does not define the lie in the same way in all sources. For a review of the Kantian notion of a lie, we suggest to see Carson, 2010, p. 71-79.

> Kant said that one should never lie. His example is that if a person who is going to kill a man asks if this one has passed this way, one should tell him the truth, even if the consequence is death. For the pedantry of not lying, let someone die. Kant demands of the categorical imperative that the maxim of every action applies to all, no matter what the circumstances.[27]

Borges' reference seems to be the text "On a Supposed Right to Lie from Philanthropy", 1797.[28] But Kant had referred to the lie even earlier. In the *Metaphysics of Morals*, Kant indeed describes lying as "the greatest violation of man's duty to himself, considered merely as a moral being".[29] And in the *Groundwork of the Metaphysics of Morals*, moreover, he uses the example of the fulfilment of promises to differentiate action out of duty from pragmatic action, determined by contingent ends.[30] We are not interested here in dwelling on the moral or legal treatment of lying,[31] which is certainly a complex subject that nevertheless exceeds the limits of this paper.[32] Rather, we are concerned here with the anthropological aspect of the lie detailed in the lecture notes of the critical decade. This aspect finds that lying is an insight into our dark representations. In particular, we are concerned to highlight that, according to the sources analysed, it seems that for Kant there is something unconscious in the foundation of our lies. Reason and its imperatives do not constitute the only motive for us to be truthful.

In the lectures the situation in physics is compared with the situation in moral research. In the first case, we are always dealing with knowledge of which we are conscious. In the case of morality, on the other hand, men's motivations are lost in a multitude of dark representations. In particular, it is suggested that the reason we regard lying as something that undermines men's honour is a dark representation that needs to be examined by philosophers.[33] Lying is always reprehensible, it is argued, not only because lying is wrong, but also because it affects the honour of men. The leitmotif suggested by the available notes seems to link an additional compelling reason not to lie that is connected to our sense of honour.

The concept of honour is mentioned several times in the anthropological annotations. In them, moreover, it is commented that honour is often approached as a "feeling". In the notes we have, we read that it is fashionable to call it "feeling" and that it is not clear how we could actually *feel* honour. It is interesting to note that, in the pre-critical period, Kant already referred to this feeling in particular.[34] In a pre-critical text, in fact, Kant mentions the feeling of honour as a criterion to explain national differences between diverse European cultures.[35] However, such a feeling functions in us as an admittedly hidden cause for liking some judgements of others about ourselves. It is also a reason why it is so important for us to be certain that such judgements are correct. At this point, it is suggested (and unfortunately not developed) that the idea of honour, understood as a dark representation, primarily affects the realm of private judgement and non-arbitrated interaction. The feeling of honour makes us care intimately that others think well of us and that they have an adequate representation of our virtues. This dark idea is the basis for a claim to a private sphere for the judgement about our value. The available notes suggest that it is a very hard task for philosophy to examine such thought and to understand the non-rational

27 Bioy Casares, 2006, p. 738.

28 A further discussion of the notion of lying in this text can be found in Wagner (1978), Stolzenberg (2010).

29 MS, AA 06:429.

30 GMS, AA 04:402.

31 On the moral problem, see Korsgaard (1996). On the legal approach, Granja Castro et al (2012).

32 An example of the complexity of the debates surrounding the problem of lying is Carson's thesis, for whom it is not the case that it is always wrong to lie, according to Kant. Cf. Carson, 2010, p. 79-88.

33 V-Anth/Mensch, AA 25: 871.

34 This reference is mentioned in the detailed explanation of the notion of *Gefühl* by Jean-Christophe Merle in the *Kant-Lexicon*. In the entry on moral feeling, moreover, the same author explains what this feeling means and how its conception changes with criticism. Cf. Willaschek et al 2015, p. 693ff.

35 Cf. *Beobachtungen*, AA 02: 249.

motivations for such a claim.[36] In the lectures on practical philosophy of the same period, in addition, the desire for honour is described as an original inclination (*ursprüngliche Neigung*), the reasons for which we cannot elucidate.[37]

In this way, a feeling of dark origins causes us to be concerned about the opinion of others about us and to avoid lying in order not to affect that image. This dark aspect of our mind is part of our psychology and affects how we conduct ourselves in society and how we represent our position in it.

5. Dark representations and philosophical research

In the lectures of the critical period, Kant finally takes up again an aspect of the topic that had already interested him earlier. It is the importance of dark representations for philosophical research. In this respect, we can observe that the philosopher develops two issues. On the one hand, he deals with the figure of Socrates, as the deliverer of ideas. On the other hand, he refers to the analytical aspect of philosophical research.[38]

Socrates guided his interlocutors in a search for knowledge they already had, without noticing it. According to his philosophical position, he was able to help us in the process of reminiscence by which we would reach the non-immediate and more certain truths. One function of the philosopher resembles that identified in Socrates and his metaphor of the midwife. Just as a midwife brings a child into the world, Socrates (and the philosopher in general) brings to light ideas that are dark and unconscious.[39] For Kant, this Socratic procedure, which allows us to elucidate prior knowledge, is the most suitable for philosophy. A large part of the task of philosophy is precisely identified with the analysis of concepts.[40] Usually, the task of the philosopher consists precisely in making conscious thoughts that were ignored. In this way, philosophy, when it enlightens us on such familiar matters as freedom, causality or duty, exhibits an aspect of our mind: in it our lighter ideas are linked to dark ones, and we are able to explore these links and bring the dark representations into clarity. Mention of this task of philosophy, which consists in going through our representations and trying to elucidate those that are not clear to us, can be found in numerous passages of Kantian work, even before criticism.

Thus, for example, in his *Inquiry Concerning the Distinctness of the Principles of Natural Theology and Morality*, a text written in 1762, he explains the task of philosophy as an analytical procedure. In this science, we are confronted with a concept that is already given, the concept of a thing. But this concept is for us confused or indeterminate. For this reason, we must analyse it and determine the notes and their relations. Kant describes a sequence of actions that is necessary in the definition of the philosophical concept. First I must decompose it, then I must compare the assembled notes with the concept. Finally, I must make the abstract thought I already have become detailed and determined. Kant illustrates his explanation by means of an example. It is the concept of time. Kant states the following: we all have a concept of time. Now, this concept needs to be elucidated. To do this, I must study its idea in all its relations, to discover notes by means of its decomposition. Then I must put the various notes together and see whether they constitute a sufficient concept. Finally, I must see if there are redundancies,

36 V-Anth/Mensch, AA 25: 869f.

37 V- Prakt Phil/ Pow, AA 27: 221.

38 For a detailed examination of the model of science Kant had in mind and how he redefines the notion of analysis, see Engfer (1982).

39 V-Anth/Mensch, AA 25: 871; V-Anth/Mrong, AA 25: 1222.

40 V-Log/ Dohna-Wund, 24: 780.

that is, if any of those notes are contained in the others.[41] A philosophical definition is the explanation of the content of an idea, i.e. of the representation of a thing. The task of the philosopher is to make our unclear representations comprehensible. The idea that we have dark or unclear representations that we need to enlighten and that it is the task of the philosopher to elucidate is recurrently developed in the sources.

Later, in the *Critique of Pure Reason*, the procedure of *exposition* with which Kant begins the investigation of sensibility seems to point to the same thing, as does the explanation of the method of philosophy that we find in the "Discipline of Pure Reason in its Dogmatic Use". However, the analytical procedure that characterises philosophical knowledge does not exhaust what the latter is.[42] Moreover, analysis is not only identified with the capacity to make our dark representations conscious. What we read in the anthropology notes is, more precisely, that the analytical procedure is nourished by the dark representations in us.

In the empirical psychology that is suggested, without being the main topic, throughout the anthropology lectures, dark representations seem to occupy an important place in understanding the mechanisms of thinking, not only in the specific field of philosophy. The dark representations present themselves as a foundation on which our conscious representations are based. For our philosopher, *discovering* something seems to be nothing more than tracing the links that bind our representations, whether they are clear or not, in order to shed light on the obscure ones. *Remembering*, similarly, is about bringing ideas out of the darkness and into the light. That is to say, remembering is to make conscious representations already available but neglected.[43]

Concluding remarks

Dark representations are unconscious representations. Kant does not specify in the sources examined what he means by this. But he does provide some metaphors that allow us to reconstruct some kind of explanation. It is about the larger part of the contents of our mind, which can be thought of as a vast map. In it, only some parts are illuminated. These clear portions are supported by the others and allow us to identify them. The procedure of identifying dark representations is a procedure that looks for the causes of certain effects.

In this article we have gone through Kant's considerations about our dark representations. In particular, we have dealt with the contents of the anthropological annotations corresponding to the critical decade. In them, as we have seen, it is still maintained that *there are* dark representations. For the philosopher, such representations would become evident from their effects, so we conclude that they exist by inference. We have examined some cases in which we are able to make such an inference. In particular, we have looked at how dark representations manifest themselves in our fears, in our feeling of respect, in analytical knowledge and in some of the usual procedures of our mind, such as discovery and memory.

41 Deut, AA 2: 277.

42 Mario Caimi (2012) has established that the analysis that makes it possible to illuminate obscure representations is the first of a series of steps, both according to the text of 1762 and according to the "*Methodenlehre*" of the *Critique of Pure Reason*.

43 V-Anth/ Mrong, AA 25: 1221.

Bibliographic References

ANDERSON, E. 'Emotions in Kant's Later Moral Philosophy: Honour and Phenomenology of Moral Value'. In: Betzler, Monika (Ed.): *Kant's Ethics of Virtue*. p. 123-145. Berlin: de Gruyter, 2008.

BAUMGARTEN, A. G. *Metaphysica*, 41757. In AA 15, 5-54 and AA 27, 5-226

BAYEFSKY, R. 'Dignity, Honour, and Human Rights: Kant's Perspective'. In: *Political Theory*, 41, 2013, p. 809–837.

BIOY CASARES, A. *Borges*. Buenos Aires: Destino, 2006.

CAIMI, M. 'Application of the Doctrine of Method in the critical examination of reason'. In: *Studia Kantiana*, 13, 2012, p. 5–16.

CARSONS, T. L. *Lying and Deception: Theory and Practice*. Oxford: Oxford University Press, 2010.

DIOGENES LAERCIO. *Vitae philosophorvm, recognovit brevique adnotatione critica instrvxit*. H. S. Long, Tomvs Prior, Oxford: Clarendon, 1966, p. 122-134.

ENGFER, H.-J. *Philosophie als Analysis*. München: Frommann-Holzboog, 1982.

GIORDANETTI, P. 'The Unconscious as Root of Kants A Priori Sentimentalism'. In: Piero Giordanetti, Riccardo Pozzo, Marco Sgarbi (Ed.) *Kant's Philosophy of the Unconscious*. p. 77-88. Berlin, Boston: De Gruyter, 2012.

GRANJA CASTRO, D. M. and CHARPENEL ELORDUY, E.. 'Kant y el escandaloso tópico de la prohibición de mentir: una lectura alternativa desde la filosofía del derecho'. In: *Endoxa*, 29, p. 15-44, 2012.

KANT, I. *Gesammelte Schriften* Ed.: Vol. 1-22 Preussische Akademie der Wissenschaften, Vol. 23 Deutsche Akademie der Wissenschaften zu Berlin, since Vol. 24 Akademie der Wissenschaften zu Göttingen. Berlin 1900ff.

KORSGAARD, C. 'The right to lie: Kant on dealing with evil'. In: *Creating the Kingdom of Ends*. p. 133-158. Cambridge: Cambridge University Press, 1996.

LOCKE, J. *An essay concerning human understanding*. P. H. Nidditch (Ed.). (The Clarendon Edition of the Works of John Locke). J. W. Yolton (Ed.) Oxford: Oxford University Press. 1975.

SÁNCHEZ RODRÍGUEZ, M. 'Pequeñas percepciones e Ilustración en Leibniz y Kant. Una revisión de la interpretación deleuziana de Leibniz'. In: *Revista Filosofía Universidad de Costa Rica*, 51, p. 281-289, 2012.

SÁNCHEZ RODRÍGUEZ, M. 'La teoría de los tipos de representación en Leibniz y sus principales influencias en la estética y la lógica de la Ilustración alemana'. In: *Cultura*, 32 p. 271-295, 2013.

SCHÖNECKER, D. 'Kant's Argument from Moral Feelings: Why Practical Reason Cannot Be

Artificial'. In: H. Kim & D. Schönecker (Ed.), *Kant and Artificial Intelligence*. p. 169-188. Berlin, Boston: De Gruyter, 2022.

SHERMAN, N. 'The Place of Emotions in Kantian Morality'. In: A. O. Rorty/O. Flanagan (eds.). *Identity, Character and Morality. Essays in Moral Psychology*. p. 149-170. Cambridge: Cambridge University Press, 1990.

STOLZENBERG, J. 'Kant und das Recht der Lüge'. In: *Kantian Journal*, 32, p. 7-16, 2010.

WAGNER, H. 'Kant gegen "ein vermeintes Recht, aus Menschenliebe zu lügen"'. In: *Kant-Studien*, 69, p. 90-96, 1978.

Recebido 11/2022 | Aprovado 12/2022

A Feminist Perspective on Kant in the Context of Social Irrationality[1]

[Uma perspectiva feminista sobre Kant no contexto da irracionalidade social]

Sandra Markewitz[1]

Universität Vechta (Vechta, Alemanha)

DOI: 10.5380/sk.v21i1.91541

Abstract

A feminist reading of Kant faces several different problems, although Carol Hay's account gives a stimulating example. Outlining the harms of oppression as a denial of equal liberty in the tradition of John Stuart Mill, it will be pointed out how Kant's thinking can support the idea that there is a feminist goal of counteracting oppression and how Kant's anthropological writings in particular serve –heterogeneous in nature – as a reminder of women's potential rational qualities, whereas on the other hand the task of "preserving the species" binds women to ensuring the persistence of the natural order. That, in the matrimonial union, one partner must "yield to the other" leads to a claim of superiority that seems outdated – inequality, today, cannot be seen as equality anymore – it clashes with the demand that we should think for ourselves using our rational capacities. In this vein two feminist objections are raised: Kant is privileging the rational over the animal and the rational over the emotional. In further discussion the question arises if Kant's work can be neutralized or if he is to be read in a way not disguising his strictness. Finally social irrationality is thematized: distraction from the goal of rationality can be perceived as utterly human.

Keywords: Feminism; Embodiment; Emotions; Irrationality; Oppression.

1 sandra.markewitz@uni-vechta.de

Carol Hay begins her account of Kantianism, Liberalism, and feminism with a rather provocative question: "What does the philosophy of a bunch of dead white men have to tell us about oppression? Rather a lot I will argue".[2] Feminism in a Kantian perspective raises several different questions: What can a feminist defense of Kant look like, are there elements in Kant's ethical and anthropological approach that do not belong in the defense model, and, which role plays the aspect of social irrationality in the feminist actualization of Kantian thought?

The harms of oppression shall be analyzed with reference to the resources of Kantianism. Oppression is defined by Hay with reference to John Stuart Mill and his *The Subjection of Women* as "a denial of equal liberty, which, for Mill is tantamount to a denial of the opportunity to develop one's rational capacities for thought and action. Mill's central concern in the *Subjection of Women* was to establish that 'the legal subordination of one sex to the other … is wrong in itself, and…one of the chief hindrances to human improvement'".[3] So the need for decreasing oppression in the interactions of the sexes is bound to a specific idea of progress. Progress as "human improvement" (Mill) denies that kind of oppression and subjection, which hinders equal liberty and its expression in human society. The question of progress and the danger of being progressistic – pursuing change for its own sake, thereby making progress an ideology – shall not be considered further. Apart from the overall question how feminism and Kantian thinking go together, there are two other points I shall address: How Kant's thinking can support the idea that there is a feminist goal of counteracting oppression and, secondly, that there is, with Hay and Kant, a "general obligation to resist oppression".[4] These questions are connected to the general focus of social irrationality given that "our capacity for practical rationality can (…) be harmed when we face illegitimate restrictions on the full and proper exercise of these capacities".[5]

1. Kant and Women: A Case of Prejudice?

Hay's provocative claim concerning dead white men telling women how to live has two sides. On the one hand, Kant is quoted with remarks such as: "Scholarly women use their books somewhat like a watch, that is, they wear the watch so it can be noticed that they have one, although it is usually broken or does not show the correct time".[6] This corresponds to the second end of nature pertaining women, that is, "the cultivation [*Cultur*] of society and its refinement by womankind".[7] Cultivation itself is a step in the development of reason in humankind. It contributes to the enlightenment, even though using books as decorative items may not look like, at first sight, as an expression of the development of reason. Their time and perception of time is specific in a distracted and distracting way, as if they could use books only as a decoration of a vain and futile endeavor. On the other hand, Kant seems to have a misogynist view of women, for instance in this quotation:

> When nature entrusted to woman's womb its dearest pledge, namely the species, in the fetus by which the race is to propagate and perpetuate itself, nature was frightened so to speak about the preservation of the species and so implanted this fear – namely fear of physical injury and timidity before similar dangers – in woman's nature; through which weakness this sex rightfully demands male protection for itself.[8]

2 Carol Hay, Kantianism, Liberalism, and Feminism. Resisting Oppression, Basingstoke, UK: Palgrave 2013, p. viii.

3 Vgl. Ebenda, p. 2. The Quotation of Mill stems from: "John Stuart Mill, The Subjection of Women", in J.S. Mill and Harriet Taylor Mill, Essays on Sex Equality, Alice S. Rossi (ed.), Chicago: University of Chicago Press 1970 (1869), p 125.

4 Ebenda, p. x.

5 Ebenda, p. 123.

6 Anth, AA 7: 307; Hay, 2013, p. 51

7 Anth, AA 7: 306.

8 Anth, AA 7: 306.

What is called biologistic today, binds women to the procreational purpose. Perpetuating the species means for women to be a container or vessel of something growing in them, not being an agent putting herself out in the world and to worldly purposes apart from the procreational one. A question to ask concerns the nature of protection, which is seemingly a necessity when facing women's weakness. But is this correct? When is the protection of women needed? Women need protection against men harming them in a specific societal and cultural context, not because of an innate shortcoming rooted in "woman's nature". Kant's anthropological writings, thus, present a limited view of women because of an unreflective biologistic stance on the one hand and a false implication of the kind of violence women need to be protected against on the other: this violence against women lies in most cases in the behavior of men, especially in the private realm, not in some anonymous threat. Thus, it is a paradox that women should seek help and protection in those who – seeing it in the binary difference of the sexes – are the one's trying to harm them. This does include personal acts of violence, abuse, or dominance and structural violence against women, namely hindering them to pursue university education or raising their voice publicly, etc. Kant's stressing of a female vulnerability rooted in procreation, in the fact that women can carry a child, seems to exclude women from practices in which autonomy is required in current society. That they should – as servants – not vote, as they are dependent on someone else (*Wes Brot ich ess, des Lied ich sing*), narrows women's possibilities of expression further.

Another quotation from Kant's anthropological writings, in *The character of the sexes*, illuminates how perceived weakness of women led to the assumption that

> in all machines that are supposed to accomplish with little power just as much as those with great power, art must be put in. Consequently, one can already assume that the provision of nature put more art into the organization of the female part than of the male; for it furnished the man with greater power than the woman in order to bring both into the most intimate physical union, which, in so far as they are nevertheless also rational beings, it orders to the end most important to it, the preservation of the species.[9]

We find here two heterogeneous elements of thought: the one considering women as rational beings, which puts Kant ahead of his time, the other insisting on women's responsibility to preserve mankind, in order to ensure the persistence of the species. The latter, in Kantian times, was nearly undisputable. It was seemingly not feasible to imagine women being as free as men, thus being able to abstain from motherhood. In other words, rationality was attributed to women, but together with attributing eternal responsibility for the fetus. That the mother should not be rational is not of Kant's concern, which is, in fact, at first glance, a rather modern thought. Motherhood and being responsible for preserving the species, indeed, do not hinder rationality. But, still today, the mother is an over determined figure in our culture. Motherhood is one of the few areas where women are allowed to exert authority over others, namely, their children. But this kind of authority is, firstly, bound to the private realm and, secondly, only allowed over children, who, in turn, are meant to take on as adults their conventional gender roles. The authority of the mother stems not from her being a rational human being but from her role in reproducing the old order in which her own place, in most cases, is only a subordinate one. We can see in this context how the ascription of rationality is also intended for women, but, at the same time, they are bound to their species in a way that oppresses them. However, not if motherhood were their real preference. But, in Kant's lifetime till today, women are often forced to want something for the sake of the persistence of the conventional order. Motherhood, in many cases, is an adaptive preference only. Relating back to the quotation above, today we would expect an ascribed rationality to show in behaviors and actions that women for a long time could not produce – they had, for instance, no access to higher education, let alone political leadership. But when women in this limited way have been considered as "rational", it was, as we see from today's point of view, often a rather empty word. Today we would expect

9 Anth, AA 7:303.

additional and more specific signs of someone being rational, i.e., the ability to decide, to choose, to judge in not merely a reactive way. Kant's prejudice against women seems to deny such capacities but does not explicitly exclude women from further development. It is after this first look at Kant's thoughts on the character of the sexes in his anthropological writings that he ascribes rationality to women as a mere potential, not something women could de facto express in self-governed actions and beliefs, because they, for the sake of the fetus and their own fragile nature, must depend on others.

The union of two rational beings, then, says Kant must follow a pattern where "one partner must *yield* to the other, and, in turn, one must be superior to the other in some way, in order to be able to rule over or govern him".[10] The problem of this claim is that superiority from one over the other is in Kant a thought based on inequality seen as equality: "each partner must be superior in a different way: the man must be superior to the woman through his physical power and courage, while the woman must be superior to the man through her natural talent for mastering his desire for her; on the other hand, in still uncivilized conditions superiority is simply on the side of the man".[11]

The different kinds of superiority are meant to make, for instance, matrimonial life equal. Can superiority be ascribed in such fundamentally different ways? It is obvious that, especially seen from a contemporary perspective, the superiority of women that Kant concedes exists merely in reference to the man. It is not rooted in the woman herself. Mastering a man's desire is, moreover, a rather fragile kind of superiority, because beauty fades over time, so the desire may not be there anymore. It is as well a weak basis for the woman's superiority to rely on, i.e., being liked or wanted by a man for physical features, as opposed to being superior because of innate qualities such as the physical power of men. Even when a man gets older and one could argue his physical power fades, in most cases, the average man will still be physically stronger than the average woman. This kind of superiority specific to women is ephemeral, and it is not appropriate to rational agents. In fact, this kind of superiority pertains only to the body, not to their mental strength, such as their capacity of imagination. The physical strength of men is a source of superiority that may hurt women. Hence, it is undesirable. The ability to awaken men's desire is useless in cases in which this desire is extinguished in a particular man. Furthermore, it supports the traditional patriarchal order and its role allocation. It is, moreover, a kind of superiority that is dependent on another person, but superiority cannot depend on the fleeting wants and needs of another person even if the matrimonial bond is strong.

What can be seen in Kant's anthropological writings is that he traces back the differences between the sexes to biological circumstances. Biologism does, as we can see today, close the possibilities of explanation when talking about men and women's roles in society. The idea of gendered role-play itself does not occur to Kant. He takes what he sees as facts. As we can conclude from the former examples, the ascription of rationality has different implications for men and women: The power of handling men's desire seems rather irrational today since it can wane at any moment. Over the centuries, it became clearer, although it is still contested today, that oppression begins not with outright acts of violence against women, but when they are not considered as victims of a whole system of oppression.[12] On the other hand, Kant concedes very clearly in *What is Enlightenment?* (1784) the oppressed status of women as *Hausvieh* which are *held am Gängelwagen*, quiet animals not daring or wanting to live without *Vormünder*, legal guardians of many a kind. There seems to be a tension in Kant's work: seeing the general oppression of women, their subordinate status in society, and, at the same time, ascribing (in his anthropological writings) rationality or potential rationality to both sexes – in so far as the expression of authority is different, which we would mostly criticize from today's point of view as a double standard. Is it not precisely the difference of states of superiority that ensures a

10 Anth, AA 7: 303.

11 Anth, AA 7: 303.

12 Hay, 2013, p. 7.

lasting inequality? And, to address an even more complicated point: Is the connection between the sexes (matrimonial or not) in need of a reciprocated condition of superiority? When this word is seen as an expression and concept with positive connotations, the union is in danger of being a place of negotiating authority more than negotiating the good life or even emancipatory side effects, such as enabling each other to develop one's own character, to do beneficial things to society, etc.

The idea of someone leading the union is meant to serve rationality – of both partners and the union between a man and a woman itself. As we might put it now, rational behavior begins with self-government not with being governed by others. However, on a broader scale, could it be a sign of freedom to resist the demand of someone close to us to lead since we should judge for ourselves as rational human beings? Furthermore, how is a woman's rationality to be combined with Kant's claim "she does not shy away from *domestic warfare*, which she conducts with her tongue and for which nature endowed her with loquacity and eloquence full of affect, which disarms the man?"[13] This description of the shared household is almost a caricature and, moreover, repeating a hackneyed cliché concerning one of women's weaknesses consisting in her talking too much. Furthermore, in this scenario, the man even "relies on the right of the stronger to give orders at home because he is supposed to protect it against external enemies; she relies on the right of the weaker to be protected by the male partner against men, and disarms him by tears of exasperation while reproaching him with his lack of generosity".[14] The right of the stronger is one classic aspect of the definition of power: to be able to force someone to do something against their own will. Forcing someone is at the heart of the union, again, as part of a protection that would not be needed if men would not behave as enemies towards women. To conclude these considerations concerning Kant's possible prejudice against women: Prejudice is already there, when superiority is considered a necessary part of marriage, be it on both sides though shown by different capacities. That women should be led is still a misunderstanding today – her alleged weakness is only weak in uncivilized circumstances. However, if the philosopher can discern the *Gängelwagen* which is holding women back from their true potential, forcing them to live under the control of legal guardians, then, seen from our point of view, this insight is not consistent. The *Gängelwagen* cannot be overcome by replicating old prejudices about women being matrimonial monsters or as the sex that is "constantly feuding with itself, whereas it remains on very good terms with the other sex".[15] This is, in fact, not a genuine character trait of all women or lack of good conduct, but deeply rooted in all women being treated, and used to being treated, as subordinate in societal entities. It must be remembered that women were and still are in many cases not allowed to exert real authority over other grown-ups, only in the context of caring for others who depend on them for a limited amount of time.

2. A feminist reading of Kant

Kant's view of women understands, on the one hand, that women are oppressed in a fundamental way by the *Gängelwagen*. This insight is, on the other hand, not followed through. Biologistic prejudice merges with knowing that women are not free. When Carol Hay states, "it is clear enough that Kant himself was no friend to women",[16] this, in my view, is to be taken with a grain of salt. She differentiates further: "I count myself among the feminist philosophers who think there is no reason to insist that these anthropological views must fully infect Kant's central philosophical views".[17] I wonder if the metaphor of an infection is fully suitable here.

13 Anth, AA 7: 304.
14 Anth, AA 7: 304.
15 Anth, AA 7: 305.
16 Hay, 2013, p. 51.
17 Hay, 2013, p. 51.

One should not separate good causes of autonomy and the value of the moral law from bad aspects of a philosophical stance. Hay admits, "the society in which Kant lived was deeply sexist; that this sexism is sometimes apparent in the works of someone writing against such background social conditions should be just as unsurprising".[18] How does this feminist reading treat important aspects of Kant's work? I would like to outline and discuss – with Hay – two criticisms feminist readers may have in reference to Kant's philosophy. Given the limited amount of time in our context, these outlines are given in a rather broad way. The first criticism "focuses on Kant's privileging the rational over the animal" and the second consists of Kant's "privileging the rational over the emotional".[19] The main aspects of these criticisms are bound to a positive view on Kant, because his philosophical account "of duties to the self" is needed to help feminist critiques to explain the problem of "the gendered norms of self-sacrifice that have historically exploited women".[20]

Can those considered as weak have duties to the self? Especially when a "self" is not ascribed to them due to sexist standards in society? Mustn't women accept the adaption of numerous values foreign to them? Can personhood and centuries of oppression go together? Can one deliberately choose to be a person (personhood) with an ideal of autonomy when still today the *Gängelwagen* traditionally holds women back? This restriction affects not only a surface of regulations in a governmental reality, but limits women's idea of themselves, if any, because the habit of socially deprecating ascriptions to women leads, in most cases, to a limited, non-imaginative, overly pragmatic view of the self and its capacities. Women learn to adapt to their circumstances in a different sense than men do.

A feminist view on Kant not only states the obvious – that women had no equal rights in Kant's lifetime – but also illuminates the implications of someone having no possibility of seeing oneself in a benign, appreciating way which is reflected by the values of a societal entity and thus strengthened.

This is directly connected with the first feminist criticism of Kant: His privileging the rational over the animal. Women have been considered as non-rational throughout the centuries: as decorum, object, or commodity to the behalf of men using them in different segments of society and among all social strata. Their being sophisticated or wealthy in most cases was not attributable to them, but rather bound to the duty of representation, of representing a man's status in society – father or husband – and sticking to social hierarchy. Apart from the fact that every human being represents something via physical impression, way of speaking, behaving, etc., the expectation of representing others, not myself, alienates women from genuine self-expression. In our societies, it is a person not being transcendent, per de Beauvoir's sense in *The Second Sex*, but only immanent (bound to a small radius of self-expression not going out into the world) that makes them not rational enough to act from moral principles. The animal is rationality's enemy. The way women are looked at directly affects their worth as moral agents; they are human beings of a lesser value, needing protection as all domesticized animals do.

Hay's premise in dealing with Kant, then, is that "Kant's work can be read in a gender-neutral manner and thus we can use his moral framework for explicitly feminist purposes".[21] I would not agree with this premise. Briefly, I shall simply say that the problem of neutralizing Kant's view is that the seemingly neutral concepts, such as autonomy and duty, could carry dangers of their own since only the ascription of rationality appears to enable human beings to do good things to others. In any case, we face a neutralized philosopher made useful against the background of a former sexist society.

But back to the feminist criticism of Kant: In the *Doctrine of Right* in the *Metaphysics of*

18 Hay, 2013, p. 51.

19 Hay, 2013, p. 50.

20 Hay, 2013, p. 50.

21 Hay, 2013, p. 51.

Morals, Kant states that women lack "civil personality".[22] They are "passive citizens", "mere underlings of the commonwealth because they have to be under the direction or protection of other individuals, and so do not possess civil independence"[23] (MM 6:314-315).

Apart from that, the so-called Kantian split affects women: We are split beings, separated into an animal and a rational side. Hay refers to Martha Nussbaum's remark on the Kantian split in her article *The Future of Feminist Liberalism* (2004). Nussbaum claims that our dignity *is* the dignity of an animal. "It is the animal sort of dignity, and that very sort of dignity could not be possessed by a being who was not mortal and vulnerable…".[24] Oppressing the animal side means to miss a huge part of what it means to be alive as we are, to carry our human endowment intertwined with the animal features. When can animality be perceived as a threat to rationality? Only, as we might put it, when the latter is overemphasized, because the good things human beings are capable of must be thought of as being rooted in the rational aspect of their lives. Could, with Nussbaum, human vulnerability (the general need for protection) be an emancipatory potential to exerting individual freedom? If so, then the difference between men and women that Kant stresses in his anthropological remarks as well as in the *Metaphysics of Morals* would wane. Nussbaum claims, too, that animality "itself could have a dignity".[25] We see, for instance, in Eisler's *Wörterbuch*, that dignity in Kant's moral system is a fixed term. Can we change meanings as we like, since the context of Kant's work determines the meaning of "dignity"? This would adapt dignity to our present ideas and circumstances but decontextualize Kant's use of the term. On a fundamental level we have to decide if conceptual aspects, such as decontextualizing Kant's central expressions, should be possible or not. This is not only true for feminist critiques of Kant, but for any discussion of philosophically or culturally problematic terms. Apart from that *Würde*, dignity for Kant is the idea that is followed by the "sittliche Handeln".[26] Does Hay's attempts to neutralize Kant's work, rendering it prolific for feminist discourse, destroy a given philosophical architecture? Or, is decontextualization serving as an unexpected enhancement of outdated philosophical claims and anthropological views?

Dignity, in the Kingdom of Ends, is antagonistic to those things having a price. So the feminist criticism that connects dignity and animality would devalue what the *Kingdom of Ends* is about. Already the first feminist criticism of Kant, that animality is separate from rationality and ascribed to women, finds reference in his anthropological writings: women's philosophy is "not to reason, but to sense".[27] This first feminist criticism is called *The embodiment objection*: our bodily existence has to be oppressed to secure the course of reason. The oppression of women as only sensing not reasoning beings is part of the more general oppression human beings should want for their own moral advancement. It seems as if soft standards, such as the vulnerability Nussbaum refers to, would alter the implication of Kant's idea greatly and lead us to the questions outlined in the first feminist criticism (the embodiment objection) as to how far a well-meant neutralization of Kant can go and if it is possible without destroying Kant's work. Feminist justice, as we have seen so far, comes with decontextualization and extracting the good from the bad where it might also be better to draw conclusions from the seemingly neutral aspects of Kant's work. Maybe the feminist criticism, as Hay presents it in the beginning of her book, should address concepts of autonomy as an androcentric male obsession[28] or universalism: "anti-colonialist critics point out that liberalism's ideals have been implemented historically in ways

22 Vgl. Hay, 2013, p. 52.

23 Quoted in Hay, 2013, p. 52.

24 Martha Nussbaum, "The Future of Feminist Liberalism", In: Baehr (ed.), Varieties of Feminist Liberalism, Lanham: Rowman and Littlefield 2004, p. 106, quoted in: Hay, Kantianism, Liberalism, and Feminism, p. 53.

25 Nussbaum, "The Future of Feminist Liberalism", p. 106, quoted in: Hay, 2013, p. 53.

26 Rudolf Eisler, Kant-Lexikon. Nachschlagewerk zu Kants Sämtlichen Schriften / Briefen und handschriftlichem Nachlass, Hildesheim: Olms 1964, S. 612.

27 Immanuel Kant, Observations on the Feeling of the Beautiful and Sublime, Frierson and Guyer (eds.), Cambridge 2011 (1764) (Cambridge University Press), p. 132-133, quoted in Hay, 2013, p. 51.

28 Vgl. Hay, 2013, p. 15.

that consistently ignore the experiences of the oppressed members of society".[29] Universalism, then, does not include women as oppressed members of society, whatever "member" may mean. Not all members of a group or larger social entity are treated equally. As long as we stick to the usual way those entities are perceived, there is an inner hierarchy of each social entity or group. Perception, in this respect, shapes experience. The embodiment objection of feminist critique reminds us of the "fundamentally embodied and social character of our rational nature", namely, that rationality cannot be split from our bodily experience. When women are described as those needing protection in Kantian anthropology, being creatures not reasoning but sensing, it seems that they belong or should belong to the animal side of existence. This devalues women in the Kantian approach that "locates every bit of our value in our rationality…".[30] The depiction of women as depending on others, furthermore, weakens their capacity of being autonomous – they stand under the law of others as servants or as underprivileged men do. The rational person is depicted as independent, capable of making choices, thinking, judging, obeying only principles that are, of course, not adaptive nor false preferences unlike the weak and dependent things women have to do as those being colonialized, e.g., creating a place for others to thrive, not deciding themselves what to do in their lives, etc. Can they live beyond adaptive preferences? Martha Nussbaum calls the idea of the rational being, as Hay points out, "the fiction of competent adulthood".[31] This means that feminist discourse, again, reminds us of our shared vulnerability. Does privileging the rational pay tribute to our actual endowment as men and women, where vulnerability, even a fragile constitution is how we begin and how we end? In sum, the first feminist criticism is that Kant splits the rational and animal in human beings although these two aspects of life are intertwined. Or, can the objection be defused by saying that Kant's claims are only normative? Yet isn't normativity deeply connected with how a descriptive approach is shaped? Does there not seem to be a gap between, for instance, saying that women need shelter and therefore can only lead in a marriage by holding a men's desire (which, to the average feminist view, is rather unattractive) and being a rational being with dignity and, according to Kant, *sittliche Handeln*? When women are closer to the animal side of things, "only" sensing, this seems to devalue them in a moral architecture. Apart from that, returning to Hay's constructive approach to Kant, she defends the philosopher by stressing how "we are fundamentally distinct from the rest of nature".[32] We are not, as animals, unpredictable, and it is precisely this element of unpredictability, uncanniness or, using a Wittgensteinian expression, that we do not know our way around with creatures other than humans, that is the threatening aspect of animality. In short, we cannot trust instinctive behaviors when our own instincts are seen not as natural or positive but as hindrances to the rational adulthood of men. Rationality being grounded in the noumenal rather than in the phenomenal realm, has, as Hay puts it in her feminist view, led to interpret Kant "as conceiving rationality as *fundamentally immune* to the sorts of harms that our animal bodies are vulnerable to".[33] Again, rationality is a normative claim as the best human beings can possibly be, but why? Is Kant's epistemological project of securing science against former mere metaphysical philosophical models tied to anthropological views devaluing what needs protection and is not autonomous? Feminist criticisms remind us that we are not only rational agents and should, maybe, not strive to be. Women are capable of rationality – as men are – but something seems lost when the principle has no exception: Is it not the fragile part making us human, our dependency, moodiness, childlike sense of wonder and intuition (*Ahnungsvermögen*)?

Since we cannot dwell on this further, let's move forward to the second feminist objection, what Hay calls the *emotions objection*. Kant is privileging "the rational over the emotional".[34] The

29 Hay, 2013, p. 27.

30 Hay, 2013, p. 53.

31 Nussbaum, "The Future of Feminist Liberalism", p. 105, quoted in Hay, 2013, p. 54.

32 Hay, 2013, p. 54.

33 Hay, 2013, p.55.

34 Hay, 2013, p. 56.

emotional aspects of our lives seem to be neglected; we are not seen fully as human beings. Hay refers to Kant talking "as if inclinations are nothing more than a hindrance to the possibility of morality".[35] She not only criticizes this from a feminist point of view but refers to Friedrich Schiller arguing in 1794 that "man can be self-opposed in a twofold manner; either as savage, if his feelings rule his principles, or as barbarian, if his principles destroy his feelings".[36] Kant also says that inclinations should not be destroyed totally, but held in their place, because, when not hindered, they grow. In any case the fiction of the competent adult having mastered his feelings could become, with Schiller and the feminist emotions objection, a cold and hard caricature of a lively human being. It is not only the general *Menschenbild* / human image this seems to imply that is problematic. A person that is morally excellent but oppressing their feelings pays a huge price for excelling morally, leading a life that is a shadow of the full human experience. This is precisely (and this is how embodiment and emotions objection are connected) what Martha Nussbaum meant when she stressed our shared dependent human nature. This shared nature should not lead to oppressing important aspects of ourselves. Because the oppression of women in ordinary societies is, since women are commonly prejudiced as being the emotional sex, to be equated with the oppression of emotions in a society where the rational is seen as the most human expression of humans. In contrast, the feminist claim that the rational should not be privileged over the emotional is what makes us human. The discussion is centered not only around the role of the emotional itself, but around an underlying concept of humanity. In paraphrasing Carol McMillan, feminist thought characterizes Kant as a philosopher saying that no action springing from natural inclination, rather than duty, can have moral worth.[37] Instead, the feminist emotions objection reminds us that "liking" or "preference" should also be taken into account when talking about what it means to act for human beings and what it means for human beings to act morally. The objection boils down to an insistence that Kant is wrong to think that morality must be grounded in something non-contingent".[38]

Contemporary Kantians, as Hay puts it, have argued for the moral framework against the emotions objection.[39] We find this with Marcia Baron, in "The Alleged Moral Repugnance of Acting from Duty", Barbara Herman, in "On the Value of Acting from the Motive of Duty", or Thomas Hill, in "The Importance of Autonomy".[40] Hay then stresses that "nothing in Kant's view implies that acting in the absence of emotions, or in opposition to them, is required or even desirable from a moral point of view".[41] But does this not mean to defuse Kant again, being unable to bear his rigor? It reminds me of people removing all challenging elements from a religion, making it into something like a feel-good attitude instead of something where obeying god is crucial. Can the feminist critique of the emotions objection be met by saying that Kant was actually not that serious about oppressing one's feelings, etc.? In the end, the objection brings to light different forms of answering the critique: The easy way out, i.e., Kant did not mean it in such a strict way, or the one I feel is more ethical in nature, that is, insisting on Kant challenging the reader and demanding a lot but sticking to it. An underlying question of the two feminist objections might be if oppressing women as emotional – following the path of prejudice - might lead not only to their own irrationality and strengthen this impression, but also to social irrationality. One solution to see women as ordinary human beings and not as something beyond the male norm is presented by feminist criticism: Valuing our rational nature

35 Hay, 2013, p. 57.

36 Schiller, 1854, p. 14.

37 Vgl. Hay, 2013, p. 57. Carol Mc Millan is quoted from her book: Carol McMillan, Women, Reason, and Nature, Princeton 1982 (Princeton University Press), p. 20.

38 Hay, 2013, p. 57.

39 Vgl. Hay, 2013, p. 58.

40 Vgl. Marcia Baron, "The Alleged Moral Repugnance of Acting from Duty", in: Journal of Philosophy 81 (1984), p. 197-220; Barbara Herman, "On the Value of Acting from the Motive of Duty", in: Philosophical Review 90 (1981), p. 359-382; Thomas Hill, "The Importance of Autonomy", in: Women and Moral Theory, ed. by Eva Feder Kittay and Diana T. Myers, Totowa: Rowman and Littlefield 1987, p. 129-138.

41 Hay, 2013, p. 58.

and the idea of self-respect.

3. Feminism and Social Irrationality

Criticisms are only hints not a cure. To say that denying women a capacity for rationality might itself be irrational, as one might imagine a kind of clever remark, is only a superficial trick. It reminds me of the argument stating that men and women should have equal rights in the workplace, because this increases productivity and general output. The process of neutralizing Kant, cleaning his work, so to speak, from its misogynist elements, also has dangers of its own. Yes, women can be rational, but should they only be rational, and should men be either? Along Hay's lines of argumentation, Kant's legacy and importance for feminist discourse consists, above all, in reminding us of two capacities of human beings that make their lives worth living: rational nature and self-respect. Whereas the oppression of will can make human beings weak and destroy their self-image as well as any ambition or ability to thrive, it can also make these oppressed persons act irrationally.[42] Yet, this is differentiated further, because "it is not the members of oppressed groups who are in danger of acting irrationally; it is the members of oppressor groups".[43] For Mill, as we heard in the beginning, oppression consists in not being able to develop one's rational capacities for thought and action. This becomes clear when self-respect is also seen as an element of the rational characteristic. Can irrational subjects respect themselves? And is not respecting oneself an acceptance of oppression? Hay stresses that self-respect consists in recognizing the intrinsic value that other people have and respect them accordingly, in the same vein, we must recognize our own intrinsic value and act accordingly.[44]

It seems to be an implicit belief that after recognizing one's worth, the capacity to act rationally will go along with it. If we remember the feminist goal to counter-act oppression, we must create contexts in which distractions from this goal, such as human exhaustion,[45] are not overlooked but accepted as utterly human. Not for the sake of productivity – this would again be a progressistic mistake – but because the animal and emotional side of humans not being machines must be considered by any concept that tries to overcome social irrationality. Self-respect, as one of Kant's legacies to feminist discourse, means being able to act for one's causes. It is a highly ambivalent goal because, as we can conclude, it is not always possible for women to express "a righteous and self-respecting concern"[46] for themselves, thus also fostering the common goal of social rationality. The oppressed had to learn the opposite: that their concerns for themselves did not matter. Showing concern for myself to prevent social irrationality on a larger scale needs a history, a practice of self-care, even pride; the kind of *Selbstverhältnis*, self-relationship, which is mirrored and strengthened in the contexts a potentially moral subject is surrounded by. Feminist objection does not only refer to the role of women, but to the traditional social role-play of men alike. Changes can come, when resisting oppression is seen as a necessity to all, not only an obligation of some.

42 Vgl. Hay, 2013, p. 127.

43 Hay, 2013, p. 127.

44 Vgl. Hay, 2013, p. 66.

45 Vgl. Hay, 2013, p. 129.

46 Bernard R Boxill, "Self-respect and Protest", in: Philosophy and Public Affairs 6 (1976), p. 58-96, 61. Quoted in Hay, 2013, p. 113.

Bibliographic References

BARON, M. 'The Alleged Moral Repugnance of Acting from Duty'. In: *Journal of Philosophy*, 81, 1984, p. 197- 220.

EISLER, R. *Kant-Lexikon. Nachschlagewerk zu Kants Sämtlichen Schriften / Briefen und handschriftlichem Nachlass*, Hildesheim: Olms, 1964.

HAY, C. *Kantianism, Liberalism, and Feminism*. Resisting Oppression, Basingstoke, 2013.

HERMAN, B. 'On the Value of Acting from the Motive of Duty'. In: *Philosophical Review*, 90, 1981, p. 359- 382.

HILL, T. 'The Importance of Autonomy'. In: Feder Kittay, E. and Myers, D. T. (eds.) *Women and Moral Theory*, Totowa: Rowman and Littlefield, 1987, p. 129-138.

ROSSI, A.S. 'John Stuart Mill, The Subjection of Women'. In: Mill, J. S. and Taylor Mill, H. *Essays on Sex Equality*. Ed. by Rossi, A.S. Chicago: University of Chicago Press, 1970.

SCHILLER, F. *The aesthetic letters, essays, and philosophical letters of Schiller*. Translated by N. Weiss, Boston: Little and Brown, 1854.

Recebido 11/2022 | Aprovado 02/2023

"The taking-something-to-be-true [that] cannot be communicated": remarks on the (lack of) communicability to understand the problem of social irrationality

["O considerar-algo-verdadeiro [que] não se pode comunicar": observações sobre a (carência da) comunicabilidade para compreender o problema da irracionalidade social]

Rômulo Eisinger Guimarães[1]

Universidade Federal do Paraná (Curitiba, Brasil)/Universität Vechta (Vechta, Alemanha)

DOI: 10.5380/sk.v21i1.91542

Abstract

Since learning how "to make use of your *own* Understanding" (WA, AA 08: 35), and "seeking the supreme touchstone of truth in oneself (i.e., in one's own Reason)" (WDO, AA 08: 146, footnote) – in a word: to enlighten oneself – is a process; and since Reason, in such a process, "needs attempts, practice and instruction" (IaG, AA 08: 19), it is only slowly that one can arrive at Enlightenment (cf. WA, AA 08: 36). As a matter of fact, the process of Enlightenment involves culture, i.e. the cultivation of the human being and his or her rational predispositions. And of fundamental importance in this process are communicability, participation and publicity – i.e., the possibility of making an idea public. Now, since these abilities are so important for Enlightenment, everything that does not allow the communication of ideas and thoughts to each and every human being is diametrically opposed to this process. This seems to be the case with the phenomenon of social irrationality and cognitive vices. The present paper aims to highlight how the problem of (lack of) communicability proves essential to understanding the phenomenon of social irrationality.

Keywords: social irrationality; vice of subreption; communicability; Enlightenment.

Resumo

Uma vez que aprender a "fazer uso de seu *próprio* Entendimento" (WA, AA 08: 35), e a "procurar em si próprio (isto é, na sua própria Razão) a suprema pedra de toque da verdade" (WDO, AA 08: 146, nota de rodapé) - em uma palavra: iluminar-se – é um processo; e uma vez que a Razão, em tal processo, "precisa de tentativas, de exercício e de aprendizagem" (IaG, AA 08: 19), é apenas lentamente que se pode chegar ao Esclarecimento (cf. WA, AA 08: 36). De fato, o processo de Esclarecimento envolve a cultura, i.e., o cultivo do ser humano e suas predisposições racionais. E de fundamental importância neste processo são a comunicabilidade, a participação e a publicidade – i.e., a possibilidade de tornar uma ideia pública. Ora, uma vez que estas habilidades são tão importantes para o Esclarecimento, tudo o que não possibilita a comunicação de ideias e pensamentos para todo e qualquer ser humano é diametralmente oposto a este processo. Este parece ser o caso do fenômeno da irracionalidade social e dos vícios cognitivos. O presente artigo tem como objetivo destacar como o problema da (falta de) comunicabilidade se mostra essencial para a compreensão do fenômeno da irracionalidade social.

Palavras-chave: irracionalidade social; vicio de subrepção; comunicabilidade; Esclarecimento.

1 r_eisinger_guimaraes@yahoo.com. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, within the scope of the project *Kantian Perspectives on social irrationality*.

1.

"The greatest" – Kant claims in the second chapter of the *Canon of pure Reason* –

> and perhaps only utility of all philosophy of pure Reason[2] is [...] only negative, namely that it does not serve for expansion, as an *organon*, but rather, as a discipline, serves for the **determination of boundaries**, and instead of discovering truth it has only the silent merit of **guarding against errors** (KrV: A 795; B 823).[3]

As a matter of fact, such a "determination of boundaries" seems to be nothing but the very critical task of Kantian philosophy. For, if one admits that "critical" derives from the verb *krínō* <κρίνω>, i.e., to separate (Bailly, 1935, p. 688), Kant's critical philosophy seeks nothing but a separation: of legitimate from illegitimate thinking, aiming to know "how much I may hope to settle [with Reason] if all the material and assistance of experience are taken away from me" (KrV: A XIV). And this attempt, of answering with pure Reason the questions of Reason itself, Kant admits, may on the one hand seem a "pretentious and immodest" claim; on the other, however, it is an enterprise which is

> incomparably more moderate than those of any author of the commonest program who pretends to prove the simple nature of the *soul* or the necessity of a first *beginning of the world* <*Weltanfang*>. For such an author pledges himself to extend human cognition beyond all bounds of possible experience (KrV, A XIV).

Now, to raise ourselves "beyond all bounds of possible experience" with our *a priori* cognition (KrV: B XXI), Kant points out elsewhere, is the wish of Metaphysics, that "battlefield of [...] endless controversies" (KrV: A VIII); that

> battlefield [...] that appears to be especially determined for testing one's powers in mock combat; on [which] no combatant has ever gained the least bit of ground, nor has any been able to base any lasting possession on his victory [for] there is no doubt that up to now the procedure of metaphysics has been a mere groping, and what is the worst, a groping among mere concepts (KrV: B XV).

The controversies of Metaphysics seem "endless" precisely because this discipline completely exceeds our human capacities. Independent of all experience, Metaphysics "is nothing but the inventory[4] of all we possess through pure Reason, ordered systematically" (KrV, A XX). Thus, it is it, the Metaphysics, the discipline that is concerned with constituting the system of pure Reason, with presenting the whole of philosophical knowledge under a unity, under a "systematic interconnection" (KrV: A 841, B 869). Kant's philosophy, in turn, is "incomparably more moderate" than any other philosophical program – so Kant himself believes – because "it is a treatise on the method, not a system of the science itself", i.e., of pure Reason, "catalog[ing] [its] entire outline [...], both in respect of its boundaries and in respect of its entire internal structure", Kant's philosophy does nothing but "the entire preliminary sketch" of such a system (KrV: B XXII).

However, one might ask, how would such a sketch be drawn? And to that Kant will answer: insofar as Reason "has insight only into what it itself produces according to its own design"; insofar as it "tak[es] the lead with principles for its judgments according to constant laws"; insofar as it "compe[ls] nature to answer its questions, rather than letting nature guide its movements by keeping Reason, as it were, in leading-strings"; finally, insofar as Reason approaches nature

2 I would like to stress at the outset that the terms "Reason", "Judgment", "Understanding", "Imagination" and "Sensibility" are capitalized whenever they refer to faculties.

3 The references to Kant's works follow the guidelines laid down by the *Akademie-Ausgabe*. As for the emphasis, *italics* refer to the original works; **bold** are my own emphasis.

4 Such an inventory would, according to Kant, be divided into four parts: ontology, rational physiology ("doctrine of nature of pure Reason"), rational cosmology, and rational theology (cf. KrV: A 845, B 873).

> with its principles in one hand, according to which alone the agreement among appearances can count as laws, and, in the other hand, the experiments thought out in accordance with these principles – yet in order to be instructed by nature not like a pupil, who has recited to him whatever the teacher wants to say, but like an appointed judge who compels witnesses to answer the questions he puts to them (KrV: B XIII).

Kant's critical philosophy is, in this sense, "the science of the relation of all cognition to the essential ends of human Reason (*teleologia rationis humanae*), and the philosopher is not an artist of Reason but the legislator of human Reason" (KrV, A 839, B 867): a legislator serving a court, that "is none other than the *Critique of pure Reason* itself" (KrV, A XII). It is, then, through the sieve of this court that the philosophy of pure Reason allows its "silent merit" to shine: to prevent us, to keep us away from errors. Errors among which we can consider the "cognitive vices".

In order to try to highlight how and in what sense the lack of communicability, publicity and participation can help in understanding the causes of social irrationality and cognitive vices, as well as, on the other hand, the culture of these skills is essential to the idea of a progress of the humankind, collaborating, therefore, in its process of Enlightenment, this paper begins by showing that, although "vice", in general, is a recurrent – and, it can be argued, plurivocal – topic in the *Corpus Kantianum*, a specific kind of vice seems directly related to the problem of social irrationality: the vice of subreption (*section 2*). Roughly describable as a "taking-the-(merely)-subjective-for-objectiv", a vice of subreption seems to fall short of that touchstone that distinguishes science from both faith and opinion: namely, the possibility of communicating and rendering a judgment, a thought or an idea true, i.e., valid not only for the subject, but for each and every rational human being. And by losing such a "connection with the truth", such a judgement also loses the criteria of communicability, publicity and participation (*section 3*). Without these criteria, however, any and every thought proves to be a problem not only for communication, in a more general way; rather, it also shows itself to be an obstacle on the path towards humankind's adulthood: its Enlightenment (*section 4*).

2.

On "vices", in general, Kant seems both prolix and inconstant: he deals with the subject in different works, but with different emphases – sometimes more, sometimes less in-depth. From pre-critical writings – such as *Dreams of a spirit-seer* (cf. TG, AA 02: 339; 372), *Universal natural history and theory of the heavens* (cf. NTH, AA 01: 329; 347; 357; 365), and *The only possible argument in support of a demonstration of the existence of God* (cf. BDG, AA 02: 104) – to the texts of his philosophical maturity – such as the three *Critiques* (cf. KrV: A 779, B 807; KpV, AA 05: 38; 88; KU, AA 05: 314; 327), the *Groundwork of the metaphysics of morals* (cf. GMS, AA 04: 442; 458) and the *Religion within the boundaries of mere Reason* (cf. RGV, AA 06: 24 ss, footnote; 26-7; 31; 33-4; 37-8; 47-8; 58; 120) – passing through the *Anthropology from a pragmatic point of view* (cf. Anth, AA 07: 149; 171; 276; 298; 301; 325) and *The metaphysics of morals*, to stay with few examples, we find some lines on this topic. It is, however, in the latter – i.e.., in *The metaphysics of morals* – where Kant seems to typify, to index, in the *Doctrine of the elements of ethics*, vices, opposing them to duties of each and every human being. It is there where Kant lists for us not only habits contrary to the duties to oneself, but also those contrary to the duties we have towards other human beings.

Attempting against the "duty of a human being to himself as an animal being" i.e., the duty of self-preservation, we have, e.g., the vices of "**murdering oneself**" <*Selbstentleibung*> (MS TL, AA 06: 422), and "**mai[ming]** <*verstümmeln*> **oneself**" (MS TL, AA 06: 423) – the former would be the vice o of "willfully *killing* oneself" (MS TL, AA 06: 422); the latter, of the deprivation "of

an integral part or organ" (MS TL, AA 06: 423). In addition to these, Kant mentions the vices of the "**defiling oneself by lust**" <*Selbstschändung*> or "lewdness" <*Unkeuschheit*> (MS TL, AA 06: 424), and "**stupefying oneself** <*Selbstbetäubung*> **by the excessive use of food or drink**" (MS TL, AA 06: 427) – respectively: "the unnatural use of [one's] *sexual inclination*" (MS TL, AA 06: 420), and the "*excessive consumptions*" (MS TL, AA 06: 420), such as "*drunkenness*" <*Versoffenheit*> and "*gluttony*" <*Gefräßigkeit*> (MS TL, AA 06: 420), which weaken "one's *capacity* for the natural (and so indirectly for the moral) *use* of one's powers" (MS TL, AA 06: 421).

Contrary to "human being's duty to himself merely as a moral being" (MS TL, AA 06: 428), i.e., "without taking his animality into consideration" but taking into account only "what is *formal* in the consistency of the maxims of his will with the *dignity* of humanity in his person" (MS TL, AA 06: 420), Kant quotes the vices of **lying** <*Lüge*>, **avarice** <*Greiz*> and **servility** <*Kriecherei*> **or false humility** <*falsche Demut*>. The first one, as a refusal "and, as it were, annihilat[ion] of [one's] human dignity", makes the human being something of "even less worth than if he were a mere thing; for a thing, because it is something real and given, has the property of being serviceable so that another can put it to some use" (MS TL, AA 06: 429); the second one would be the restriction of "*one's own* enjoyment of the means to good living so narrowly as to leave one's own true needs unsatisfied" (MS TL, AA 06: 432) – which can be understood as a narrow-mindedness concerning the development of one's dispositions as a rational being –; and, the third one refers the denial of one's own self-esteem and moral dignity (cf. MS TL, AA 06: 435), the disregard of the human being "as a *person*, that is, as the subject of a morally practical Reason [that] exalt[s him or her] above any price" (MS TL, AA 06: 434).

Opposed to the duty of love <*Liebespflicht*> towards other human beings, we have, e.g., according to Kant, the vices of misanthropy <*Menschenhasse*> (MS TL, AA 06: 458): **envy** <*Neid*>, **ingratitude** <*Undankbarkeit*> and **joy at the misfortune of others** <*Schadenfreude*>. These concern, in turn and respectively, the "propensity to view the well-being of others with distress" (MS TL, AA 06: 458); the undermining or devaluation of "those who [...] have preceded us in conferring benefits" (MS TL, AA 06: 459); and, in the face of "the misfortune of others or their downfall in scandal", an immediate rejoicing "in the existence of such *enormities* destroying what is best in the world as a whole, and so also to wish for them to happen, is secretly to hate human beings" (MS TL, AA 06: 460).

Finally, with regard to the "duties of virtue toward other human beings arising from the respect due them" (MS TL, AA 06: 462), Kant points out the vices of **arrogance** <*Hochmut*>, **defamation** <*Afterreden*>, and **ridicule** <*Verhöhnung*>: the first one would be a "demand that others think little of themselves in comparison with [us]", "a solicitation on the part of one seeking honour for followers, whom he thinks he is entitled to treat with contempt" (MS TL, AA 06: 465); the second one – also called by Kant "backbiting" <*übele Nachrede*> – concerns an "the immediate inclination, with no particular aim in view, to bring into the open something prejudicial to respect for others" (MS TL, AA 06: 466); and the third one relates to the "*wanton faultfinding* and *mockery*, the propensity to expose others to laughter, to make their faults the immediate object of one's amusement [...] in order to deprive [them] of the respect [they] deserves" (MS TL, AA 06: 467).

Now, Kant offers us this whole inventory of vices, of this "brood of dispositions opposing the law", of these "monsters [one] has to fight" (MS TL, AA 06: 405) in the practical sphere – particularly, in the *Doctrine of Virtue*: it is there that he tells us, in a very broad way, that "it is when an intentional transgression [of a duty] has become a principle that it is properly called a *vice* (*vitium*)" (MS TL, AA 06: 390). It is, however, in the *Doctrine of Right* where he tells us of another vicious behaviour that, I believe, comes closest to the aforementioned cognitive vices.

This is because it is also a kind of "vice", Kant considers,

> a common fault <*ein gewöhnlicher Fehler der Erschleichung*> (*vitium subreptionis*) of experts on right to *misrepresent*, as if it were also the objective principle of what is

> right in itself, that rightful principle which a court is authorized and indeed bound to adopt for its own use (hence for a subjective purpose) in order to pronounce and judge what belongs to each as his right, although the latter is very different from the former (MS RL, AA 06: 297).

Such a **"taking-the-(merely)-subjective-for-objective"** is an error, a fault that Kant points out in different areas. In his moral philosophy, for instance, he tells us of "an optical illusion in the self-consciousness" with respect to what determines the will in moral action. In what sense? Kant recognises that "moral disposition is necessarily connected with consciousness of the determination of the will *directly by the* [moral] *law*". And, he follows, "consciousness of a determination of the faculty of desire <*Begehrungsvermögen*> is always the ground of a satisfaction in the action produced by it" (KpV, AA 05: 116) – a satisfaction, Kant seems to suggest elsewhere, in the "represented connection of the subject" i.e., of his or her will, with "that which is good absolutely and in all respects, namely the morally good" (KU, AA 05: 209).

As a matter of fact, "since this determination has exactly the same inward effect [...] of an impulse <*Wirkung eines Antriebs*> to activity, as a feeling of the agreeableness <*Annehmlichkeit*> expected from the desired action would have produced", Kant recognizes, "we easily look upon what we ourselves do as something that we merely passively feel [that satisfaction, that agreeableness] and take the moral incentive <*die moralische Triebfeder*> for a sensible impulse <*sinnlicher Antrieb*>" (KpV, AA 05: 116-7). However, he points out, such a satisfaction arising from the awareness of the determination "is not the determining ground of the action: instead" – he follows – "the determination of the will directly by Reason alone is the ground of the feeling of pleasure, and this remains a pure practical, not aesthetic [i.e. sensible], determination of the faculty of desire" (KpV, AA 05: 116). To consider the contrary, i.e., that it is satisfaction that is the determining ground of action, would be to "demea[n] and defor[m] the real and genuine incentive, the law itself – as it were, by means of a false foil – by such spurious praise of the moral determining ground as incentive as would base it on feelings of particular joys (which are nevertheless only results)" (KpV, AA 05: 117).

If so, that "optical illusion in the self-consciousness" to which Kant was drawing our attention does not seem to be anything other than this: the illusion of mistaking the moral motive for the sensible impulse; of taking a subjective satisfaction – and not the objective law itself – as the determining ground of moral action. And to hold what is done driven by sensible impulse as the principle, as the determining ground of moral action Kant calls a "*vitium subreptionis praticum*" (HN, AA 15: 454) – i.e., a practical vice of subreption.

Such a vice of subreption entails, then, providing false reasons, false grounds in order to justify something.[5] But this confusion between subjective and objective can take place, one can argue in two ways: if one takes the sensible for the intelligible, or if one takes the intelligible for the sensible. In the first case, this means taking sensible impressions, of the senses, i.e., sensations for a thought of objective validity; in the second, taking merely subjective thought as empirically and objectively valid.

Now, to take the impression of the senses – or sensations – as an objective thought, as an objective and therefore universally valid principle, is the ethical problem mentioned above – namely: that to remain in the field of sensations, of feelings undermines the understanding of morality itself, i.e. its rational basis. As a matter of fact, if morality revolved around the conservation of the subjects, their satisfaction, their welfare (cf. GMS, AA 04: 365), "[they] needs and inclinations, the entire satisfaction of which [they] su[m] up under the name happiness" (GMS, AA 04: 405), to all this "an implanted natural instinct would have led much more certainly" (GMS, AA 04: 396) than Reason.

However, to ascribe to morality the principle of happiness is, for Kant, rather problematic: for well-being is not ruled by to a good conduct; for it is entirely different to make a person happy

5 Kant will say elsewhere that a vice of subreption "gives the Understanding a false direction" (EEKU, AA 20: 222).

and to make them good; and, above all, because, in this way, one "bases morality on incentives that undermine it and destroy all its sublimity" (GMS, AA 04: 442) – and such a sublimity follows from the fact that morality has to be valid for all rational beings, "and *only because of this* be also a law for all human wills. On the other hand, what is derived from the special natural constitution of humanity – what is derived from certain feelings and propensities [...] – that can indeed yield a maxim for us but not a law"; "the sublimity and inner dignity of the [moral] command" – Kant goes on lines below –"is all the more manifest the fewer are the subjective causes" (GMS, AA 04: 425).[6]

On the other hand, if one takes the intelligible for the sensible, placing the "Understanding before appearance", "reflection before intuition", Kant will say, "one thus deprives the Understanding of its rights" (HN, AA 15: 92). In which sense? In the sense that the proof of the objective reality of a concept [through which we think something] depends, precisely, on its reference to intuitions (cf. KU, AA 05: 351).

If one of the three questions on which the whole interest of Reason dwells is "*what can I know?*", Kant flatters himself that he has "exhausted all possible replies to it, and finally found that with which Reason must certainly satisfy itself and with which, if it does not look to the practical, it also has cause to be content" (KrV: A 805, B 833), namely: I can know what enters my sensible intuition through the pure forms of Space and Time and is adequate to the pure concepts of the Understanding through the transcendental Judgment. This is the entire legitimate theoretical use of Reason in forming empirical cognition of phenomena; a use with which "we might perchance be content" or should have to do so by "necessity, if there is no other ground on which we could build [legitimate cognition]" (KrV: A 236, B 295), than in the domain of pure Understanding. And by "domain" I mean here that which, referring to a cognitive faculty, dictates the principles to its objects (KU, AA 05: 174). Thus, the domain of the Understanding is the legislation by means of concepts of (or in relation to) nature – as object of the senses (cf. KU, AA 05: 195).

To place thought **before** (or, perhaps, **above**, or even **apart from**) appearances deprives the former of all its rights because, one might think, it subtracts from thought all its legitimate field of action. This is the very reason why Kant considers that *Transcendental Logic* corresponds to a "Logic of truth": precisely because it deals with the legitimate use of what we have to build empirical cognition on what is sensibly intuited, occupying itself "with laws of the Understanding and Reason, [...] solely insofar as they are related to objects *a priori*" (KrV: A 57, B 81). In addition, not surprisingly, the first part of that Transcendental Logic goes from "the guideline <*Leitfaden*> to the discovery of all pure concepts of the Understanding" (KrV: A 66, B 91) to the conclusion that "we cannot *think* any object except through categories [,and, furthermore,] we cannot *cognize* any object that is thought except through intuitions that correspond to those concepts" (KrV: B 165).

This is, as a matter of fact, the difference that Kant establishes between **cognizing** merely **thinking**. For "cognition", he says, requires two elements:

> first, the **concept**, through which an object is thought at all (the category), and second, the **intuition**, through which it is given [, so that] if an intuition corresponding to the concept could not be given at all, then it would be a thought as far as its form is concerned, but without any object, and by its means no cognition of anything at all would be possible, since, as far as I would know <*wüßte*>, nothing would be given

6 Similar is what can be observed, for instance, in aesthetics: if one keeps the judgement within the framework of sensory sensations, of what delights <*vergnügt*> or not the senses (KU, AA 05: 207), the satisfaction connected to such a judgement is therefore a pathologically conditioned one (KU, AA 05: 209), i.e., conditioned by the rejoicing or aversion that the object provokes in one's senses. Indeed, it does not seem to make sense to call "agreeable" or "disagreeable" something that does not affect my sensations positively or negatively. But, if so, the "considering-something-agreeable" or not rests on a private feeling, and here the Latin maxim seems to prevail: *de gustibus non est disputandum* – for the validity of the delight is limited solely and simply to the one who judges.

> nor could be given to which my thought could be applied[7] [...] [, for the] thinking of an object in general through a pure concept of the Understanding can become cognition only insofar as this concept is related to objects of the senses (KrV: B 146).

In short, if there is no intuition to "objectify", i.e., to give objective reality to a concept, the latter "contains only the subjective of thought, that is, the copula of the predicate in relation to the subject" (HN, AA 20: 349). And, in this case, this thought, this merely subjective concept, without intuitions, does not count as cognition.[8]

Now, what seems to be common to both vicious behaviours, i.e., to both vices of subreption, of providing spurious grounds in order to justify judgments is that, in both cases, one remains in the subjective realm (whether from a sensible point of view, or from an intellectual point of view). And, as a consequence, the judgments, arising from such biased *modi operandi*, seem to be the case of private judgments: judgments that are not subjective, but – and this seems important to be stressed – *merely* subjective.

What I am calling here "merely subjective" judgments are those that are valid only and exclusively for the one who judges. And, in this case, they differ from both judgements of objective validity and judgements of intersubjective validity. Of the first case are the judgments of logical, determined validity: judgments that rest on concepts and criteria valid for each and every person. Such are both the judgments of cognition – which rest on concepts and intuitions – and moral judgments – since the moral law, for Kant, has an objective ground, as a "*factum*" of Reason (KpV, AA 05: 31), and therefore the judgements, which are based on this law, are objectively determined, consequently, of objective validity. Of the second case, in turn, are the judgments of aesthetic validity: a subjective common validity <*Gemeingültigkeit*> (KU, AA 05: 214), which does not refer to the object and its concept "considered in its entire logical sphere, and yet it extends it over *the whole sphere of those who judge*" (KU, AA 05: 215). Such judgements could claim universal validity because they would be "sanctioned by transcendental grounds" (Zammito, 1992, p. 113), i.e., by the formal constellation of our a priori cognitive faculties. Of this kind are, according to Kant, aesthetic judgements-of-taste about the beautiful: judgements about "that which, without concepts, is represented as the object of a *universal* satisfaction" (KU, AA 05: 211).

In turn, private judgments <*Privaturteil*> are those that are valid solely for the subject that judges, rejecting both objective and intersubjective validity, and thus making any arguing impossible.[9] It is noteworthy, however, that Kant draws a distinction between "to argue" <*Streiten*> and "to dispute" <*Disputieren*> on some about something as follows:

> [they] are certainly alike in this, that they try to bring about unanimity in judgments through their mutual opposition, but they differ in that the latter hopes to accomplish this in accordance with determinate concepts as grounds of proofs, and so assumes objective concepts as grounds of the judgment (KU, AA 05: 338).

7 It would therefore be an empty thought (cf. KrV: A 51, B 75), i.e., without objective reality.

8 This is the reason why, it may be argued, Kant will assert that "all mathematical concepts are not by themselves cognitions, except insofar as one presupposes that there are things that can be presented to us only in accordance with the form of that pure sensible intuition" (KrV: B 147).

9 Taking advantage of the aforementioned example (cf. footnote 6), let us take into consideration, once again, the judgement about the agreeable. If, as said above, "agreeable" is what satisfies **my** inclinations, **my** preferences, not only does it not seem to make sense for someone to want to argue about it, in order to say that what I call pleasant is not so, i.e., that something does not satisfy my preferences, even if I judge it as such; rather, **it does not seem to make sense either for me to want to impute to others the validity of this judgement**, because it is a judgement that concerns only my private satisfaction. As a matter of fact, if a certain wine, a certain food, a certain colour, a certain sound delights someone, one should bear in mind that it is agreeable to him or her, whereas for another it may not be the case. To quarrel, Kant points out, about "the judgment of another that is different from our own in such a matter, with the aim of condemning it as incorrect, as if it were logically opposed to our own" (KU, AA 05: 212)., would be foolish. This is because the judgement about the agreeable, although it is an aesthetic judgement (cf. KU, AA 05: 21), involves a pleasure of merely private validity, so that "with regard to the agreeable, **the principle *everyone has his own* taste (of the senses) is valid**" (KU, AA 05: 212) and, for this very reason, it does not seem worth (or does not even seem to make sense) wasting time, i.e., arguing about it.

The absence of concepts, then, makes dispute impossible, but not arguing;[10] the absence of a common subjective validity, on the other hand, renders both disputing and arguing impossible. Now, what makes dispute possible is a logical basis, based on concepts; what, in turn, makes discussion possible seems a – between inverted commas – "minor" requirement: that something be valid beyond the very subject that judges, i.e., that his or her judgment can be, at some level, communicated. As a matter of fact, this would be, according to Kant, the "touchstone" (KrV: A 820, B 848) for distinguishing science from faith and opinion.

3.

Kant states in *Section Three of the Canon of pure Reason* that, concerning the validity of a judgement, the "taking-something-to-be-true" <*das Furwahrhalten*>, although it "may on objective grounds, [it] also requires subjective causes in the mind of [the one] who judges" (KrV: A 820, B 848). And this subjective assent, in turn, may concern either the particular nature of the subject, or a judgment of subjective validity – but not in a solipsistic way. In the first case, it would be a **persuasion**, i.e., nothing but "a mere semblance, since the ground of the judgment, which lies solely in the subject, is held to be objective". Hence, he follows, "such a judgment also has only private validity, and the taking-something-to-be-true **cannot be communicated**" (KrV: A 820, B 848); in the second case, however, it would be a **conviction**, and it may refer to a subjectivity **common to all beings endowed with healthy Reason** (which would confer objective validity on it).

Nevertheless, seen solely as an appearances of the mind of the one who judges, persuasion and conviction do not differ (cf. KrV: A 821, B 849). The difference between them arises, precisely, in the possibility of **universalizing** this "subjective", i.e., "the possibility of **communicating** it and finding it to be valid for the Reason of every human being to take it to be true" (KrV: A 820, B 848), verifying whether it concerns something that is not restricted to the subject.

Now, if, on the one hand, Kant states that the "the subjective validity of judgment [...] in relation to conviction (which at the same time is valid objectively)" (KrV: A 822, B 850) has three degrees (namely: **opining** <*Meinen*>, – as subjectively and objectively insufficient assent; **believing** <*Glauben*> – as objectively insufficient but subjectively sufficient assent; and **knowing** <*Wissen*> – as sufficient assent both subjectively and objectively), on the other hand, he seems to consider that every thought must have "a connection with truth" (KrV: A 822, B 850) if it is not to consist of a pure play of the Imagination. But what can this maintenance with the truth mean also there, where, precisely, we do not have an objective principle?

In order to answer this question, perhaps we should appeal to what is presented to us in *The Discipline of Pure Reason* with respect to hypotheses. For there Kant claims that "if the Imagination is not simply *to enthuse* but is, under the strict oversight [i.e., guidance] of Reason, *to invent*, something must always first be **fully certain and not invented**, or a mere opinion, and that is the possibility of the object itself" (KrV: A 770, B 798). And "possibility of the object itself", Kant states elsewhere, is nothing but the "general rules of unity in the synthesis of appearances, whose objective reality, as necessary conditions, can always be shown in experience, indeed in its possibility" (KrV: A 157, B 196). This means, then, that it is a matter of **guiding**

10 Which is why Kant will say that "it is possible to argue about taste (but not to dispute)" (KU, AA 05: 338). For although there is no determinate concept involved, there is, supposedly, an objectively valid ground of determination – namely: "the subjective conformity to ends <*Zweckmäßigkeit*> in the representation of an object without any end (objective or subjective), consequently the mere form of the conformity to end in the representation through which an object is given to us, insofar as we are conscious of it" (KU, AA 05: 221), i.e., "the feeling (of inner sense) of that unison in the play of the powers of the mind [the free and harmonious play of the Imagination and the Understanding (KU, AA 05: 217ss.)], insofar as they can only be sensed" (KU, AA 05: 228).

one's thinking according to the logical forms of pure thought in general, which (even without a proper application to phenomena) "contain the grounds of the possibility of all experience in general from the side of the Understanding" (KrV: B 167).

It might be interesting, in all this, to note that Kant stresses the need to start from something known in order to "anchor" thought, preventing it from straying completely from the realm of what has theoretical dignity and succumbing to daydreaming.[11] As a matter of fact, he states that even when simply opining – i.e., in a judgement according to insufficient grounds both objectively and subjectively –,

> I must never undertake to have an opinion **without at least knowing something** by means of which the in itself merely problematic judgment acquires a **connection with truth** which, although it is not complete, is nevertheless **more than an arbitrary invention**.[12] Furthermore, the law of such a connection must be certain. For if in regard to this too I have nothing but opinion, then it is all only a game of Imagination without the least relation to truth (KrV: A 822, B 850).

It is such a "*least relation to truth*" that seems to make a judgement fit to be communicated – or, at least, communicable –, i.e. fit to be valid beyond the person who utters it. And precisely this, I believe, is a problem in the case of cognitive vices.

What I would refer to, here, as a "cognitive vice" – in the context of social irrationality –, concerns an unreflective group-thinking, which involves, among other factors, suggestion and/or collective contagion. It would be a (social) phenomenon in which a group of individuals share a given (subjective) conviction, taking it for granted (i.e. as something objective). And it is in this kind of "vice" that, it could be argued, some modern problems are rooted – such as negationism, mysticism/pseudoscience, post-truth, etc.

Now, the question revolves (or seems to revolve) around something that, at first, does not seem to have an objective ground – i.e., it would not have an interpersonal basis –, valid beyond the individual subject. It seems to be, rather, something that orbits within the realm of beliefs, opinions, convictions, etc.

By itself, that would render the phenomenon of social irrationality something particularly curious. And this is because, on the one hand, as "social", it involves two or more individuals sharing the same thought; on the other hand, however, this same "shared thinking" does not seem something fit to be shared, i.e. shareable: for it lacks (or seems to lack) the criteria of "**communicability**" <*Mitteilbarkeit*>, "**participation**" <*Teilnehmung*> and "**publicity**" <*Öffentlichkeit, Publizität*>[13]. In this sense, it could be argued, social irrationality is, to a great

11 Kant also argues in terms of a (subjective) need <*Bedürfnis*> to "anchor" thought – although not in the same context, namely: in an exposition on the right of argumentation based on such a need of Reason to adapt itself "on its cognitions of supersensible objects" (WDO, AA 08: 134) – in the essay *What does it mean to orient oneself in thinking?*. There he states that Reason does not, unless it allows itself to fall into fantasies, see itself as devoid of Laws – even if these are the most general Laws of Understanding. As a matter of fact, "if Reason will not subject itself to the laws it gives itself, it has to bow under the yoke of laws given by another; for without any law, nothing – not even nonsense – can play its game for long" (WDO, AA 08: 145). This is because the final implication of the extreme freedom of thought, Kant follows lines below, when it "proceed[s] in independence of the laws of Reason", is nothing other than to "destro[y] itself" (WDO, AA 08: 146). So that "the unavoidable consequence of *declared* lawlessness in thinking (of a liberation from the limitations of Reason) is that the freedom to think will ultimately be forfeited", and whoever initially delights in the absence of legislation falls into a delusional "enthusiasm" <*Schwärmerei*> (WDO, AA 08: 145).

12 With this, Kant seems not only to argue for a certain hierarchy among convictions (*knowing*>*believing*>*opining*); he also seems to suggest a supposed fourth order of that "taking-something-to-be-true", on which he does not dwell because, one could be argued, it has no theoretical dignity, since it is something even **less than an opinion** – namely: "an arbitrary invention".

13 Although both "Öffentlichkeit" and "Publizität" could easily be translated as "publicity", there is a significant difference between these terms. While the former seems to refer more broadly to a "*bringing-into-light*", a "*making-visible/public*" – applied to a wider scope –, the latter seems to be restricted to a practical realm. As a matter of fact, "[Publizität] is about action, and more specifically political action! It is not the activity of bringing a carrot into the light of day that would have the quality of "Publizität", but the making visible of maxims and deeds in the field of

extent, an "unsocial" phenomenon.

This is because there is a deficit, if not a deliberate unconcern to make certain ideas universally valid. One thus remains in the realm of belief and opinion, short of science; in the realm of conviction, short of certainty. Nevertheless, what provokes no little surprise is the fact that, within a "microcosms" these convictions are indeed shared, communicated and even disseminated by and among its members.

In the light of this, the attempt to understand the phenomenon of social irrationality, I believe, involves explaining (or trying to explain) how cognitive vices are able of being communicated. However, answering such a question – i.e., how beliefs, opinions and intentions can be shared between different individuals – may not be so simple. This is because if, on the one hand, the *modus operandi*, i.e., the mode of reasoning of these individuals does not comply with the maxims of "the *common human Understanding* [i.e., of the "common sense (*sensus communis*)"]" (KU, AA 05: 293); on the other hand, one might argue, it could be a case of what Kant calls "positive Unreason" <*positive Unvernunft*> (Anth, AA, 07: 216).

By saying that this is reasoning that does not comply with the rules of common human Understanding, I am taking into account what, according to Kant, "as merely healthy (not yet cultivated) Understanding, is regarded as the least that can be expected from anyone who lays claim to the name of a human being" and whose possession "is certainly not an advantage or an honor" (KU, AA 05: 293); that whose maxims are "to think for oneself" (i.e., in a *unprejudiced way*); "to think in the position of everyone else" (in a *broad-minded way*); and "always to think in accord with oneself" (in a *consistent way*) (KU, AA 05: 294)[14].

On the other hand, it could be said to be a positive Unreason because, rather than being an "absence of rules <*Regellosigkeit*>" (HN, AA 16: 780), it is the use of "another rule, a totally different standpoint", faraway "from the *Sensorio communi*", in which, nevertheless, there is a "principle of unity [...] so that the faculty of thought does not remain idle" – a principle which,"although it is not working objectively toward true cognition of things, it is still at work subjectively" (Anth, AA 07: 216).

Whatever the case may be, it is because it seems rooted in prejudices and because it does not take

> account (*a priori*) of everyone else's way of representing in though, in order *as it were* to hold its judgment up to human Reason as a whole and thereby avoid the illusion which, from subjective private conditions that could easily be held to be objective, would have a detrimental influence on the judgment (KU, AA 05: 293),

that the phenomenon of social irrationality seems to run into the problem of the communicability of ideas. And as such, this phenomenon presents itself as a hindrance to the Kantian critical project.

---

political action. "Publzität" is a *political* buzzword before which even the heads of the people bend down in their state wigs <*ein* politischen *Schlagwort, vor dem in ihren Staatsperücken sich selbst des Volken Häupter bücken*> (Blesenkemper, 1987, pp. 30-1). Thus, while "Öffentlichkeit" may refer to the ability to communicate, i.e. to make public ideas, propositions, maxims, norms of behaviour, etc. – which also applies, it can be argued, to judgements-of-taste (cf. KU, AA 05: 214) –, "Publizität" seems to have a juridical-political share. This becomes clear in Kant if, for instance, one takes into account that, for him, the "*transcendental formula* of public Right" is that "all actions relating to the rights of others are wrong if their maxim is incompatible with publicity" (ZeF, AA 08: 381), and, furthermore, that "all maxims which *need* publicity (in order not to fail in their end) harmonize with **right and politics combined**" (ZeF, AA 08: 388).

14 And such maxims, which can be put, respectively, in terms of "an action independent of the reflecting Judgement", "an elevation <*Erhöhung*> above oneself as an expansion of consciousness <*Erweiterung des Bewußtsein*>", and "an identification between the general point of view and one's own way of thinking" (Kaulbach, 1984, p. 77 ss), "*in nuce* [...] contain nothing less than the normative pragmatics of our Reason, that is, a piece of the theory of Enlightenment" (Kohler, 2008, p. 145). For "the operation of reflection [...] makes it possible to activate the common sense mode of thinking and feeling appropriate to the broad horizon of humanity" (Kaulbach, 1984, p. 143).

4.

The cultivation, i.e., the culture of communicability, participation and publicity of ideas is, one could argue, fundamental in Immanuel Kant's philosophy. Understood both as a process and as the result of a process of "production of the aptitude of a rational being for any ends in general (thus those of his freedom)" (KU, AA 05: 431), the concept of "culture" plays a key role in what might be called the Kantian project of Enlightenment.

If by "production of the aptitude of a rational being" is meant the "to cultivate (*cultura* <*Anbau*>) his natural powers" (MS TL, AA 06: 444), "culture" can be understood as "cultivation of oneself", i.e., as self-development of the rational being. And for such a development, Kant points out, two principles must be obeyed: "'live in conformity with nature' (*naturae convenienter vive*)" and "'*make yourself more perfect* than mere nature has made you' (*perfice te utfinem, perfice te ut medium*)" (MS TL, AA 06: 419).

Now, to live in conformity with nature implies living according to the ultimate end that nature has reserved for the human race (cf. KU, AA 05: 431). But what does nature have in store for us as beings endowed with Reason? Since nature has been very parsimonious with human beings (giving them "neither the horns of the steer, nor the claws of the lion, nor the teeth of the dog" (IaG, AA 08: 19) to ensure their own well-being), it, nature, does not seem very receptive to our happiness (cf. KU, AA 05: 430). Indeed, if it were so primed for our happiness and well-being, why would it bother to give us anything other than instincts? For, as said above, a natural instinct would seem far more more efficient for this (cf. GMS, AA 04: 396).

However, if, in addition to instincts, nature has endowed us with Reason, and if, moreover, "all natural predispositions of creature are determined sometime to develop themselves completely and purposively <*zweckmäβig*>", so that "an organ that is not to be used, an arrangement that does not attain to its end, is a contradiction in the teleological doctrine of nature" (IaG, AA 08: 18),[15] since the human being "owes it to himself (as a rational being) not to leave idle and, as it were, rusting away the natural predispositions and capacities that his Reason can some day use" (MS TL, AA 06: 444), he or she has a duty towards himself or herself to develop, to cultivate his or her Reason: this is how the human being becomes "equal to the end of his existence" (MS TL, 06: 446); it is by means of this that the human being, "as an animal endowed with the *capacity of Reason* (*animal rationabile*), can make out of himself a *rational animal* (*animal rationale*) (Anth, AA 07: 321); and this, above all, is what is meant by that "production of the aptitude of a rational being", i. e., the cultivation of the dispositions proper to a being endowed with the faculty of Reason.

In this context, to become "*more perftct* than mere nature has made you" seems to imply the recognition of this "duty to cultivate", and, furthermore, a duty to cultivate Reason as "faculty that can extend itself beyond the limits within which all animals are held" (MaM, AA 08: 112). And such an "extension beyond", in turn, means "nothing other than the transition from the crudity of a merely animal creature into humanity, from the go-cart <*Gängelwagen*> of instinct to the guidance of Reason – in a word, from the guardianship <*Vormundschaft*> of nature into the condition of freedom <*Stand der Freiheit*>" (MaM, AA 08: 115). And it is in such – and only in such! – a "state of freedom" (freedom from the fetters of animality, and freedom to make use of oneself, i.e., of one's rationality) that the human being can reach, as it were, his or her "adulthood" (cf. WA, AA 08: 35; MaM, AA 08: 116), i.e., **enlighten** himself or herself.

Yet, as the march of culture, of cultivation, of the development of the human being and his or her rational dispositions, the process of Enlightenment is a lengthy one: it does not take place in the individual, but in the species; not in a human being, but in the human**kind**.

15 In a word: if nature were to provide something that serves no purpose, that is not to be used and developed, it would at best be a sort of "spendthrift" nature.

As a matter of fact, if "in the human being (as the only rational creature on earth), those predispositions whose goal is the use of his Reason were to develop completely [...] in the individual", the latter "would have to live exceedingly long in order to learn how he is to make a complete use of [...] such predispositions": for Reason "does not operate instinctively, but rather needs attempts, practice and instruction in order gradually to progress from one stage of insight to another". Since, however, this is not the case, i.e., since "nature has only set the term of his life as short", it "perhaps needs an immense series of generations, each of which transmits its Enlightenment to the next, in order finally to propel its germs in our species to that stage of development which is completely suited to its aim" (IaG, AA 08: 19).

In the light of this, the progress of the human species, as a whole, implies the transmission of knowledge from one generation to the next one: a kind of testament, of cultural-rational heritage that one generation leaves to the following one. And for such a heredity of knowledge, the aforementioned concepts of "communicability", "participation" and "publicity" – as the "method of Enlightenment" (Habermas, 1990, p. 180), i.e., as the ability to express an idea freely and in a public way – play a fundamental role in this process.[16]

Now, if, for the progress of the human species, i.e., for its Enlightenment, the ability to communicate, to participate, and to make public one's thoughts and judgments is essential, whatever works against these abilities, one might argue, works against the process and

16 That such concepts would be essential to the idea of the progress of humanity, this Kant makes quite explicit in his writings on history, politics and right. When inquiring about a possible indicative, as "a historical sign (*signum rememorativum, demonstrativum, prognosticon*)", that could "demonstrat[e] the tendency of the human race viewed in its entirety" (SF, AA 04: 84), Kant points out that "it is simply the mode of thinking of the spectators which reveals itself **publicly** in this game of great revolutions, and manifests such a universal yet disinterested sympathy for the players on one side against those on the other, even at the risk that this partiality could become very disadvantageous for them if discovered" (SF, AA 07: 85), which proves such a moral tendency of the human race. In the case of the French Revolution, for instance, since it "finds in the hearts of all spectators (who are not engaged in this game themselves) a wishful **participation** that borders closely on enthusiasm the very expression of which is fraught with danger [,] this sympathy, therefore, can have no other cause than a moral predisposition in the human race" (SF, AA 07: 85). As a matter of fact, it is not the revolution of the barricades itself – which, as resistance by violence (on the part of the people) to violence (of a tyrannical government), is always unjust and reprehensible (cf. TP, AA 08: 299; SF, AA 07: 86, footnote) – that would point to an indication of such a progress. Rather, it is the revolution – or, more accurately, the **evolution** (cf. SF, AA 077: 87) – of thought, i.e., the free adherence to and enthusiasm for the idea it represents (the idea of a republican right), which would indicate it. In addition to this, since revolution by force is not legitimate, for effective progress it would also be necessary for the citizen to "have, with the approval of the ruler himself, the authorization to make known publicly his opinions about what it is in the ruler's arrangements that seems to him to be a wrong against the commonwealth <*gegen das gemeine Wesen*>" (TP, AA 08: 304). Indeed, since, even with a view to progress, resisting government by violence is not an option, Kant wonders: "how else [...] could the government get the knowledge it requires for its own essential purpose than by **letting the spirit of freedom, so worthy of respect in its origin and in its effects, express itself?**" (TP, AA 08: 305). Now, if it is the purpose, i.e., if it is the role of government to rule for the sake of "the commonwealth's flourishing, which is required to secure its strength and stability both internally and against external enemies, not in order, as it were, to make the people happy against its will but only to make it exist as a commonwealth" (TP, AA 08: 298-9), and if, however, this same government may "err or be ignorant of something" (TP, AA 08: 304) in the exercise of its function, it can (and must!) have its course rectified by the people! For, Kant states, "*what a people cannot decree for itself, a legislator also cannot decree for a people*" (TP, AA 08: 304). As a matter of fact, if a ruler is to represent the will of the people, then he or she must know what this will is, i.e., he or she must inform himself or herself (and be able to be informed) about the needs of the people. And, in turn, "this can happen by no other means than that of **publicity** in the event that an entire people cares to bring forward its grievances (*gravamen*)" (SF, AA 08: 89). In other words, if the government must make its people progress, make them prosper, and if, for that, it is important that the people make themselves heard, publicity, i.e., the possibility of making ideas public, seems indispensable to the march of progress, i.e., of Enlightenment. For if, Kant is quite explicit, the "*Enlightenment of the people* is the public instruction of the people in its duties and rights vis-a-vis the state to which they belong" (SF, AA 0: 89), the freedom to publicly express ideas, to make a public use of one's Reason is a *conditio sine qua non* of the process of Enlightenment of the human race (cf. WA, AA 08:26 ss): it is through it that the people become "co-legislative <*mitgesetzgebend*>" (SF, AA 07: 86, footnote) of a government; it is thus that the state can, "reform itself from time to time and, attempting evolution instead of revolution, progress perpetually toward the better" (SF, AA 07: 93). Now, to such a capacity of expressing ideas is given the name of publicity: the possibility of making "general and public judgments", due to "a natural calling of humanity to communicate with one another <*einander mitzuteilen*>, especially in what concerns people generally" <*vornehmlich in dem, was den Menschen überhaupt angeht*>. (TP, AA 05: 305). Publicity, both in Right and in Politics, would then refer to the possibility of confessing publicly (cf. ZeF, AA 08: 381), i.e., of making public an idea, a thought.

Enlightenment itself.

As a matter of fact, under the aegis of Reason, the promotion of such abilities entails the promotion of what the natural drive <*Trieb*> of each and every person claims (or, at least, should claim) – i.e., the promotion of "*sociability*", as something "necessary for human beings as creatures destined for society, and thus as a property belonging to *humanity*" (KU, AA 05: 296-7). On the other hand, the crisis of – or, more precisely, the attempts to discredit – technical-scientific knowledge, as well as its much desired supplementation by unproven ideas and theories not only (seem to) dialogue with concepts such as "social irrationality" and "cognitive vices", but also seem to oppose all that communicability, publicity and participation promote – or, at least, seek to promote.

If so, such a vicious behaviour stands as a drawback to the culture of what Kant calls "*humaniora*", i.e., of the ability to "communicate universally" (HN, AA 16: 155) with the ability "to cultivate humanity" (Refl, AA 15:604). For "*humanity* <Humanität>", he claims elsewhere, "means on the one hand the universal *feeling of participation* <*Teilnehmungsgefühl*> and on the other hand the capacity for being able to *communicate* one's inmost self universally" – properties which, "taken together [,] constitute the sociability <*Geselligkeit*> that is appropriate to humankind <*Menschheit*>, by means of which it distinguishes itself from the limitation of animals" (KU, AA 05: 355) and enforces the distinctiveness of the human being as a rational being.

Finally, it seems important – though perhaps redundant – to highlight that "**fake news**", "**self-deception**", "**pseudoscience**", "**obscurantism**" and "**post-truth**" – to stay with few examples – are quite current concepts and that orbit the phenomenon of social irrationality and cognitive vices. In order to present, from Kant's philosophy, an analysis and a problematization of the sources of these issues – and, if possible, to point out their eventual solutions – it seems indispensable, therefore, to address **the problem of the communicability of ideas**. And this in order to verify how its lack is not only possible to fit in, but is also important to understand the causes of social irrationality, frustrating a supposed progress of humankind, i.e., that ideal of Enlightenment of the human species, so advocated by Kant.

By doing so, one can humbly aim for a new understanding of Kantian philosophy – not only of its practical, social, political and juridical philosophy, but also of its theoretical philosophy, i.e. of that the so-called "court" (KrV, A XII) of Reason, approaching the broad phenomenon of social irrationality from the "necessary condition of the universal communicability of our cognition, which is assumed in every logic and every principle of cognitions that is not skeptical" (KU, AA 05: 239); by doing so, one can aim at shedding light on the not always recognized actuality of Immanuel Kant's critical philosophy – a philosophy that has, yes, its limitations and obsolescence, but keeps (at least, it was intended to show) its relevance in face of contemporary problems.

Bibliographic References

BAILLY, A. *Dictionnaire Grec–Français*. Paris: Hachette, 1935.

BLESENKEMPER, K. *„Public age"*. Studien zum Öffentlichkeitsbegriff bei Kant (*Pommersfeldener Beiträge*, Sonderbd. 4). Frankfurt a/M: Haag + Herchen, 1987.

HABERMAS, J. *Strukturwandel der Öffentlichkeit*. Untersuchungen zu einer Kategorie der bürgerlichen Gesellschaft. Frankfurt a/M: Suhrkamp Verlag, 1990.

KANT, I. 'An answer to the question: what is enlightenment?'. Transl. by Allen W. Wood. In: GREGOR, M. J. (Ed.), *Practical philosophy* (The Cambridge Edition of the Works of Immanuel Kant) pp. 353-604). Cambridge: Cambridge University Press, 1996, p. 11-22.

KANT, I. 'Anthropology from a pragmatic point of view'. Transl. by Robert B. Louden. In.: ZÖLLER, G.; LOUDEN, R. B. (Eds.), *Anthropology, history and education* (The Cambridge Edition of the Works of Immanuel Kant). Cambridge: Cambridge University Press, 2007, p. 227-429.

KANT, I. 'Conjectural beginning of human history'. Transl. by Allen W. Wood. In.: ZÖLLER, G.; LOUDEN, R. B. (Eds.), *Anthropology, history and education* (The Cambridge Edition of the Works of Immanuel Kant). Cambridge: Cambridge University Press, 2007, p. 160-175.

KANT, I. 'Critique of practical reason'. Transl. by Allen W. Wood. In: GREGOR, M. J. (Ed.), *Practical philosophy* (The Cambridge Edition of the Works of Immanuel Kant) pp. 353-604). Cambridge: Cambridge University Press, 1996, p. 133-272.

KANT, I. *Critique of pure reason*. Transl. and Ed. by Paul Guyer and Allen W. Wood (The Cambridge Edition of the Works of Immanuel Kant). Cambridge: Cambridge University Press, 1998.

KANT, I. *Critique of the power of judgment*. Transl. and Ed. by Paul Guyer; Transl. by Eric Matthews (The Cambridge Edition of the Works of Immanuel Kant). Cambridge: Cambridge University Press, 2000.

KANT, I. 'Dreams of a spirit-seer elucidated by dreams of metaphysics'. Transl. by David Walford and Ralf Meerbote. In.: D. WALFORD, D; MEERBOTE, R. (Ed.), *Theoretical philosophy*, 1755–1770 (The Cambridge Edition of the Works of Immanuel Kant). Cambridge: Cambridge University Press, 1992, p. 301-304.

KANT, I. 'First Introduction to the critique of the power of judgment'. Transl. by Paul Guyer and Eric Matthews. In.: GUYER, P. (Ed.), *Critique of the power of judgment* (The Cambridge Edition of the Works of Immanuel Kant). Cambridge: Cambridge University Press, 2000, p. 1-52.

KANT, I. *Gesammelte Schriften*. (Akademieausgabe). Berlin: Preussische Akademie der Wissenschaften / Walter de Gruyter, 1990.

KANT, I. 'Groundwork of The metaphysics of morals'. Transl. by Allen W. Wood. In: GREGOR, M. J. (Ed.), *Practical philosophy* (The Cambridge Edition of the Works of Immanuel Kant) pp. 353-604). Cambridge: Cambridge University Press, 1996, p. 37-108.

KANT, I. 'Idea for a universal history with a cosmopolitan aim'. Transl. by Allen W. Wood. In.: ZÖLLER, G.; LOUDEN, R. B. (Eds.), *Anthropology, history and education* (The Cambridge Edition of the Works of Immanuel Kant). Cambridge: Cambridge University Press, 2007, p. 107-120.

KANT, I. *Notes and fragments*. Transl. and Ed. by Paul Guyer; Transl. by Curtis Bowman and Frederick Rauscher (The Cambridge Edition of the Works of Immanuel Kant). Cambridge: Cambridge University Press, 2005.

KANT, I. 'On the common saying: that may be correct in theory, but it is of no use in practice'. Transl. by Allen W. Wood. In: GREGOR, M. J. (Ed.), *Practical philosophy* (The Cambridge

Edition of the Works of Immanuel Kant) pp. 353-604). Cambridge: Cambridge University Press, 1996, p. 273-310.

KANT, I. 'Religion within the boundaries of mere reason'. Transl. by George Di Giovanni. In.: WOOD, A. W; DI GIOVANNI, G. (Ed.), *Religion and rational theology* (The Cambridge Edition of the Works of Immanuel Kant). Cambridge: Cambridge University Press, 1996, p. 39-216.

KANT, I. 'The conflict of the faculties. Transl. by Mary J. Gregor and Robert Anchor'. In.: WOOD, A. W; DI GIOVANNI, G. (Ed.), *Religion and rational theology* (The Cambridge Edition of the Works of Immanuel Kant). Cambridge: Cambridge University Press, 1996, p. 233-328.

KANT, I. 'The metaphysics of morals'. Transl. by Allen W. Wood. In: GREGOR, M. J. (Ed.), *Practical philosophy* (The Cambridge Edition of the Works of Immanuel Kant) pp. 353-604). Cambridge: Cambridge University Press, 1996, p. 353-604.

KANT, I. 'The only possible argument in support of a demonstration of the existence of God'. Transl. by David Walford and Ralf Meerbote. In.: D. WALFORD, D; MEERBOTE, R. (Ed.), *Theoretical philosophy*, 1755–1770 (The Cambridge Edition of the Works of Immanuel Kant). Cambridge: Cambridge University Press, 1992, p. 107-110.

KANT, I. 'Toward perpetual peace'. Transl. by Allen W. Wood. In: GREGOR, M. J. (Ed.), *Practical philosophy* (The Cambridge Edition of the Works of Immanuel Kant) pp. 353-604). Cambridge: Cambridge University Press, 1996, p. 311-352.

KANT, I. 'Universal natural history and theory of the heavens or essay on the constitution and the mechanical origin of the whole universe according to Newtonian principles'. Transl. by Olaf Reinhardt. In: WATKINS, E. (Ed.), *Natural science* (The Cambridge Edition of the Works of Immanuel Kant) pp. 353-604). Cambridge: Cambridge University Press, 1996, p. 182-308.

KANT, I. What does it mean to orient oneself in thinking?'. Transl. by Allen W. Wood. In.: WOOD, A. W; DI GIOVANNI, G. (Ed.), *Religion and rational theology* (The Cambridge Edition of the Works of Immanuel Kant). Cambridge: Cambridge University Press, 1996, p. 1-18.

KAULBACH, F. *Ästhetische Welterkenntnis bei Kant*. Würzburg: Königshausen/Neumann, 1984.

KOHLER, G. „Gemeinsinn oder: über das Güte am Schönen. Von der Geschmackslehre zur Teleologie (§§ 39-42)". In: HÖFFE, O. (Hrsg.) *Immanuel Kant*. Kritik der Urteilskraft. Berlin: Akademie Verlag, 2008, p. 137-150.

Recebido 11/2022 | Aprovado 12/2022

Dieter Henrich's Early Approach to the Transcendental Deduction of the Categories[1]

[A abordagem inicial de Dieter Henrich para a dedução transcendental das categorias]

Frank Rettweiler[2]

Universität Vechta (Vechta, Alemanha)

DOI: 10.5380/sk.v21i1.91543

Abstract

The anxious over-reaction against psychologism hast still shaped Dieter Henrich's early interpretation of the proof-structure of the transcendental deduction of the second edition of the Critique of Pure Reason. From this over-reaction against psychologism, the role of the concept of synthesis and the role it plays in the transcendental deduction of categories is underestimated. This underestimation of the concept of synthesis obscures the proof structure of the transcendental deduction.

Keywords: Dieter Henrich; Kant's transcendental Deduction of Categories; Synthesis; transcendental psychology.

1 I have to thank Wayne Dornbusch for checking the text for gross linguistic errors. I am of course responsible for any remaining errors.

2 frank.rettweiler@posteo.de

Some Introductory Remarks

The German philosopher Dieter Henrich died on December 17, 2022 at the age of 95. He was one of the most important experts on the philosophy of Immanuel Kant and German Idealism. As early as the 1950s, he had written essays on Kant's philosophy, with which he opposed Martin Heidegger's interpretation of the philosophy of Kant. Later he wrote important and influential essays on the transcendental deduction of the categories of Kant's *Critique of Pure Reason*. Finally, he devoted himself more to German Idealism and wrote extensive books about its origins. In addition to his interests in the history of philosophy, he also had interests of a more systematic nature. Contrary to the *zeitgeist*, he stuck to a theory of subjectivity, which he opposed to the philosophy of some Heidegger adepts, to French postmodernist philosophy, and to the naturalistic tendencies in the Anglo-American analytic tradition. Nevertheless, he was also open to the new possibilities opening up within the analytic tradition. He had been in fruitful exchange with Peter F. Strawson, Donald Davidson, Hilary Putnam, and Héctor-Neri Castañeda.

However, we will see that he was also a child of his time and did not always oppose the *zeitgeist*. His paper on the proof-structure of the transcendental deduction of the second edition of Kant's *Critique of Pure Reason* (henceforth B-Deduction) (Henrich, 1969) was shaped by assumptions, informed by the over-anxious reaction against psychologism, that have not been questioned much since Neo-Kantianism.

Henrich's Question about the Proof-Structure of the B-Deduction

Béatrice Longuenesse has written in her landmark book *Kant and the Capacity to Judge*:

> Cohen's epistemological reading, Heidegger's phenomenological reading, and Strawson's analysis of 'transcendental arguments' have one thing in common, as paradoxical as such an agreement may seem: they all stand firmly under the banner of antipsychologism (Longuenesse, 1998, p. 5).

Dieter Henrich's interpretation of the B-Deduction in the 1969 essay about the proof-structure of the B-Deduction and also that in his book *Identität und Objektivität* (Henrich, 1976) can be added to this list. The Neo-Kantians in Marburg, as well as those of the Heidelberg and Freiburg schools (*südwestdeutscher or badischer* Neo-Kantianism, above all Wilhelm Windelband and Heinrich Rickert), already propagated a strict anti-psychologism. This anti-psychologism was reinforced by the writings of Gottlob Frege and Edmund Husserl, such that psychologism could be set aside as a failed project. No one wanted to be associated with psychologism anymore, regardless of which philosophical school they belonged to. However, when the Neo-Kantians pronounced the motto "back to Kant", they were referring to a Kant freed from all varieties and traces of psychologism. Because the Kantian psychology of capacities (*Vermögenspsychologie*) was considered a variant of psychologism, Peter F. Strawson wrote in his otherwise highly meritorious book *The Bounds of Sense*:

> I have treated the Deduction as an *argument*, which proceeds by analysis of the concept of experience in general to the conclusion that a certain objectivity and a certain unity are necessary conditions of the possibility of experience. And such an argument it is. But it is also an essay in the imaginary subject of transcendental psychology (Strawson, 1989, p. 20f).

Kant distinguished empirical from rational psychology, but (as far as I know) he did not know of a transcendental psychology. He opined that the only sentence of rational psychology is the sentence "I think". After 1785, he no longer thought much of empirical psychology, which he would have preferred to replace with anthropology. "Transcendental psychology" is a

pejorative term designed to render entire passages from the *Critique of Pure Reason* redundant and worthless for a rational reconstruction of an otherwise ingenious approach.

But Kant was very aware of the difference between psychological and epistemological, or logical, questions. It was also clear to him that epistemological or logical questions cannot be answered by means of psychology. The following quotes from §13 of the *Critique of Pure Reason* prove this:

> Jurists, when they speak of entitlements and claims, distinguish in a legal matter between the questions about what is lawful (*quid juris*) and that which concern the fact (*quid facti*). And since they demand proof of both, they call the first, that which is to establish the entitlement or the legal claim, the **deduction** (Kant, 1998, p. 219f [KrV B116][3]).

> I therefore call the explanation of the way in which concepts can relate to objects *a priori* their **transcendental deduction**, and distinguish this from the **empirical** deduction, which shows how a concept is acquired through experience and reflection on it, and therefore concerns not the lawfulness but the fact from which the possession has arisen (Kant, 1998, p. 220 [KrV B117]).

> Such a tracing from the first endavors of our power of cognition to ascend from individual perceptions to general concepts is without doubt of great utility, and the famous Locke is to be thanked for having first opened the way for this. Yet a **deduction** of the pure *a priori* concepts can never be achieved in this way; it does not lie down this path at all, for in regard to their future use, which should be entirely independent of experience, an entirely different birth certificate than that of an ancestry from experiences must be produced. I will therefore call this attempted physiological derivation, which cannot properly be called a deduction at all because it concerns a *quaestio facti*, the explanation of the **possession** of a pure cognition. It is therefore clear that only a transcendental and never an empirical deduction of them can be given, and that in regard to pure *a priori* concepts empirical deductions are nothing but idle attempts, which can occupy only those who have not grasped the entirely distinctive nature of these cognitions (Kant, 1998, p. 221 [KrV B118f]).

We cannot attest psychologism to any thinker writing this. Of course, Hermann Cohen, Martin Heidegger, Peter F. Strawson, and Dieter Henrich would agree and yet they assumed that Kant could not free himself from a certain residue of psychologism. This could be due to the fact that Kant repeatedly spoke of the capacity (*Vermögen*) of understanding or reason. Sensuality is also treated as such a capacity. Such ways of talking about these capacities could have given rise to attributing to Kant a remnant of psychologism.

But Kant's distinction between understanding and sensuality is not based on a psychological investigation. This distinction is based much more on epistemological investigations. The German philosopher Friedrich Anton Koch has further developed arguments that can already be found in Kant, for example, in the chapter "On the amphiboly of concepts of reflection" of the *Critique of Pure Reason* and also in Strawson's *Individuals*, in which he shows that numerical and qualitative identity diverge (Strawson, 1971, p. 121ff). If we assume, following Koch, that qualitative identity is based on purely conceptual descriptions and we further assume that there can be two different water drops that do not differ in terms of conceptual description, then numerical and qualitative identity are not identical. However, on a level of purely conceptual description, numerical and qualitative identity cannot be distinguished. If we now ascribe to ourselves the ability to describe things as well as to distinguish things in space and time, despite fitting the same description, then we could also be credited with the capacity of having descriptive concepts as well as the capacity of identifying particulars in space and time. The first capacity is what Kant calls "understanding" and the second is "sensibility". The distinction is based on epistemological and logical considerations, but not on psychological ones. So if we attribute a psychology to Kant, then it should be an epistemological psychology, not a transcendental one. Kant's ostensibly psychological considerations thus turn out to be epistemological or logical

3 KrV: so I will give the original pagination of the Critique of Pure Reason 1787.

considerations. However, antipsychologism suspected that whenever capacities were discussed, a psychological argument crept in where an epistemological or logical argument actually belonged. If Kant was then to be freed from the alleged psychologism, what remained was an argumentation framework that was of course only a ruin compared to the structure that Kant originally presented to us. This is what happens to the argumentation within the transcendental deduction of the categories, if one considers the synthesis as an eliminable psychological remainder. Strawson in particular viewed synthesis as part of a discarded transcendental psychology, but Henrich paid no attention to the synthesis either.

It is striking that Henrich pays almost no attention to §15 of the B Deduction, in which synthesis plays a prominent role. Instead, Henrich does not let the argument of the transcendental deduction begin until §16, where the original-synthetic unity of apperception is presented as an unmediated premise. For Strawson, too, this original-synthetic unity of apperception appears as a premise, which he then also interprets as ordinary empirical self-consciousness. But Kant is not only concerned here with the self-ascription of thoughts, which, however, can be overlooked if one excludes the topic of synthesis from the outset. Henrich also overlooks the fact that the original-synthetic unity of apperception as a theme is already prepared in §15, although one can hardly imagine that such a gifted and masterful Kant scholar isn't aware of this. This can only be explained by the fact that the subject of synthesis is to be consciously avoided.

A similar observation can be made regarding the question "from the way in which the manifold of an empirical intuition is given" (Kant, 1998, p. 253 [KrV B144]). The German word "wie" can be translated as "how", but is in the Cambridge Edition translated as "the way in which". I only mention this because I want to call this question the "question of how". Since Neo-Kantianism, this "question of how" has been dismissed as a psychological one. If the topic of synthesis is to be avoided and the "question of how", then not much that makes sense can be gained from the second proof step in the B-Deduction.

If one avoids the "question of how", then the question arises as to what contribution the second proof step in the B-Deduction adds at all. In 1969, Henrich apparently advocated the so-called restriction thesis more out of embarrassment than out of complete conviction. This restriction thesis stated that in the second proof step of the B-Deduction a restriction had to be lifted. But what restriction should be lifted? At the end of the first paragraph of §21, which is the transition from the first to the second proof step Kant writes (emphasis mine):

> In the sequel (§26) it will be shown from the **way in which** the empirical intuition is given in sensibility that its unity can be none other than the one the category prescribes to the manifold of a given intuition in general according to the preceding §20; thus by the explanation of its *a priori* validity in regard to **all objects of our senses** the aim of the deduction will first be fully attained (Kant, 1998, p. 253 [KrV B144f]).

Kant is here talking about "all objects of our senses". But that doesn't mean that in the first step of the proof it's just a matter of some objects of our senses. And here, too, Kant speaks again of the way in which ("question of how") the manifold is given in intuition and its unity. If we insist, as Henrich did at the Kant conference in Marburg in 1981, that we should ignore the "question of how" (Tuschling, 1984) and look at the last sentence of §20 (Kant, 1998, p. 252 [KrV B143])[4], then there is no need for a second proof step. If I correctly understood Henrich's argumentation, as he presented it at the Kant Conference in Marburg in 1981, then, in his opinion, the decisive step in the B-Deduction had already been taken in §16. Kant opens §17 with the following two sentences:

> The supreme principle of the possibility of all intuition in relation to sensibility was, according to the Transcendental Aesthetic that all the manifold of sensibility stand under the formal conditions of space and time. The supreme principle of all intuition in relation to the understanding is that all the manifold of intuition stand

4 "Thus the manifold in a given intuition also stands necessarily stands under categories".

> under conditions of the original synthetic unity of apperception (Kant, 1998, p. 248 [KrV B136]).

The first principle was proved in the Transcendental Aesthetic, and the second principle in §16. It seems to me that going from the results in §16 to the conclusion in §20 seems to have been trivial for Henrich, so that in §16 the crucial point in the proof of the B-Deduction has already been proved. I remember that §16 always played such a prominent role in his seminars in Munich (Germany) on the transcendental deduction of the categories. But if we look more closely at the two opening sentences of §17, we can see, that Kant gives only two formal conditions for the manifold given in an intuition. And these formal conditions are only necessary but not sufficient conditions for knowledge. In the further argumentation steps after §16, Kant still has a lot to prove within the B-Deduction.

Henrich's new approach: the legal explanatory model

As I have already mentioned, Henrich never seemed to be really satisfied with the explanatory model that he presented in 1968. However, it seems to me that the question he developed in his 1969 essay is still relevant. I mean the question that Henrich raised and which Alison Laywine now calls "Henrich's Challenge" (Laywine, 2020, p.209), about the relationship between the two proof steps in a proof in the B-Deduction. According to Henrich, however, this question could only be raised if one understood the word deduction in the sense of a syllogistic deduction or a demonstration. But as we have already seen, in §13 Kant distinguishes different concepts of a deduction. An empirical deduction aims to show the circumstances under which we acquired a concept, whereas a transcendental deduction should show how a priori concepts can relate to objects. A transcendental deduction is not concerned with the history and circumstances of the acquisition of a concept, but with the justification for its use and to what extent that use is justified.

But in Henrich's view, the B-Deduction has been read far too often as a chain of syllogisms. However, if one orients oneself to §20, then it makes sense to read the B-Deduction as a proof with a polysyllogistic structure.

At the 1981 conference in Marburg, Henrich outlined a new interpretation of the transcendental deduction of the categories, which he then elaborated on in an essay from 1989. Henrich seems to have shed his great fear of psychologism a little, because he now speaks bluntly about cognitive capacities. Henrich had discovered that:

> By the end of the fourteenth century, there had come into being a type of publication that by the beginning of the eighteenth century (when it had come to widespread use) was known as *Deduktionsschriften* ('deduction writings'). Their aim was to justify controversial legal claims between the numerous rulers of the independent territories, city republics, ond other constituents of the Holy Roman Empire ... Before the final decisions of one of the Imperial Courts (which were by no means always respected), legal proceedings also required that a deduction had to be submitted by both parties (Henrich, 1989, p.32).

Because these deduction writings, like the transcendental deduction of the categories, should also justify a claim, Henrich assumed that these deduction writings contained a clue on how to read the transcendental deduction of the categories.

However, one should not assume that Henrich doubted that he regarded the transcendental deduction of the categories as a proof because of the legal background of the concept of a deduction that he had discovered. He even concedes that chains of syllogisms are present in the transcendental deduction, but doubts that considering it as a chain of syllogisms captures the specific character of the deduction. On the contrary, he rightly claims that deduction is

compatible with the most diverse forms of argument (Heinrich, 1989, p.39).

I remain skeptical whether this knowledge about the origin of the deduction writings will help us much in our understanding of the internal structure of the transcendental deduction. The version of the transcendental deduction in the first edition of the *Critique of Pure Reason* differs significantly from that in the second edition. Although the internal structure of the transcendental deduction is quite different in the two editions, both refer back to the same legal model. One could know that the transcendental deduction is a question concerning the justification of the use of concepts even without the origin of the concept of a deduction from the deductive writings. I would not go so far as to entirely deny the usefulness of this knowledge of the historical origins of the concept of deduction, but this knowledge is not sufficient for understanding the structure of the argument of the transcendental deduction in both editions. This historical knowledge cannot even explain to us why Kant completely rewrote the transcendental deduction of the categories for the second edition.

Lessons to be drawn from Henrich's and Strawson's Interpretation of Kant's Transcendental Deduction

Today the literature on the transcendental deduction has become so vast that it is difficult to keep track of it. A useful bibliography can be found in a volume edited by Giuseppe Motta, Dennis Schulting, and Udo Thiel, *Kant's Transcendental Deduction and the Theory of Apperception. New Interpretation. Berlin* (Boston 2022). Furthermore, much can be learned from the new literature on the transcendental deduction of the categories. However, I would like to conclude by offering a few thoughts on the lessons to be learned from Henrich and Strawson's interpretations.

We may assume that Kant was constantly aware of the difference between questions about the justification of the use of a concept and questions about a psychological history of the acquisition of concepts. We should not hastily accuse Kant of psychologism. Yet, that is exactly what we are doing when we dismiss much of the transcendental deduction of the categories as obscure transcendental psychology. Rather, we should try to discover the epistemological reasons for the distinctions Kant has made in the area of the psychology of capacities (*Vermögenspsychologie*). It may seem as if I am imputing the simplest of errors to such eminently important scholars as Peter F. Strawson and Dieter Henrich, but there must be an explanation for the hasty dismissal of what I have called the "question of how" and the concept of synthesis. I have no other explanation than that Kant was accused of a certain psychologism.

In his book *Kant und das Problem der Gesetzmäßigkeit der Natur* (1991), Bernard Thöle has given an explanation as to why the "question of how" plays such an important role. He recalls the main purpose of the *Critique of Pure Reason* and explains that "it lies in a critical investigation of the possibility of metaphysical knowledge" (Thöle, 1991, p. 9). For Kant, as Thöle states, metaphysics was first and foremost knowledge based merely on concepts (Thöle, 1991, p. 10). Thus, metaphysics seems to have a seductive proximity to pure science and mathematics. He goes on to characterize the main purpose of the *Critique of Pure Reason* by pointing out that Kant, in wanting to create peace in the eternal battlefield of metaphysics, therefore had to reject the epistemological claims of transcendent metaphysics. However, it must be said why these claims to knowledge are unjustified. As such, Kant has to explain the reasons why the knowledge claims of mathematics and pure natural science are justified, even though they are also based on knowledge a priori (Thöle,1991, p. 13f). It must therefore become clear why the metaphysician cannot refer to the mathematician or the pure scientist, in order to justify his own metaphysical claims with reference to the mathematician and the pure scientist. All three, the metaphysician, the mathematician, and the pure scientist, ultimately invoke the possibility

of synthetic a priori judgments. To clarify this, Thöle says: "It is particularly important that the elements of intellectual knowledge are fully explored" (Thöle,1991, p. 15). From this, it becomes clear why the "question of how" is relevant within the transcendental deduction, since, as Thöle puts it: "from the explanation of the way in which concepts a priori can relate to objects, it should result to what extent a legitimate use can be made of them" (Thöle, 1991, p. 15f).

In his 1989 essay, Henrich even argues in a similar direction when he distinguishes between an investigation and a reflection (Henrich, 1989, p. 42ff). Although the *Critique of Pure Reason* is an investigation, it must be based on reflection, since every investigation is based on reflection. Knowledge, he says, only comes about when the various cognitive faculties interlock, but only in such a way that they remain within their limits and are kept under control with the help of reflection.

> For that purpose, the mind must implicitly know what is specific to each of its particular activities. This implies, furthermore, that the principles upon which an activity is founded must be known by contrast with the other activities. Reflection consists in precisely this knowledge (Henrich, 1989, p. 42).

The principles important to the critique of pure reason, along with space and time, are the original unity of apperception. Henrich treats the original unity of apperception as the fact to which a deduction has to refer. It is, moreover, the same fact on which the reflection that forms the basis of the investigations within the *Critique of Pure Reason* is based. It seems to me, once again, in Henrich's interpretation, Kant's transcendental deduction of the categories appears too close to Fichte's foundation chapter in his *Wissenschaftslehre* of 1794/95.

Ian Proops has objected to Henrich that it is the metaphysical deduction that provides the fact on which the proof of the transcendental deduction is based (Proops, 2003)[5]. In this way, the transcendental deduction of the categories is brought back into relation to the metaphysical deduction, as Béatrice Longuenesse also strives for, as will be outlined again below.

We must not see the transcendental deduction of the categories only in the context of the large macrostructure of the *Critique of Pure Reason*, but also understand their role in the smaller structure of the Transcendental Analytic. We need to understand its role in relation to its neighboring doctrines, namely, the relation of the metaphysical deduction of the categories to their transcendental deduction and the relation of the transcendental deduction to the chapter on the schematism of the categories. Béatrice Longuenesse examined this in an exemplary manner in her book *Kant and the Capacity to Judge* (1998).

While Strawson's fear of the specter of psychologism is even more pronounced than Henrich's, Henrich has tended to view the transcendental deduction too isolated from the other tenets of the *Critique of Pure Reason*. Both the fear of the specter of psychologism and the tendency to read the transcendental deduction too isolated from other parts of the *Critique* lead to the fact that the second proof step in the transcendental deduction can no longer be understood.

In order to clarify the relationships of the transcendental deduction to the metaphysical deduction of the categories and to the chapter about the Principles of the Pure Understanding, it seems to me useful to follow the reflections of Béatrice Longuenesse from her book on Kant. There she puts forth an opposing argument concerning psychologism and the denial of the usefulness of the metaphysical deduction in Kant:

> [N]either the argument of the Transcendental Deduction of the Categories, that is, the demonstration of the role of the pure concepts of the understanding in any representation of an object, nor the System of Principles of the Pure Understanding, can be understood unless they are related, down to the minutest details of their proofs, to the role that Kant assigns to the logical forms of our judgements, and to

5 I owe the reference to this article to Nicholas Lawrence.

> the manner, in which he establishes the table of categories or pure concepts of the understanding according to the 'guiding thread' of these logical forms (Longuenesse, 1998, p.5).

With her approach to understanding the *Critique of Pure Reason*, she also opens up a way of understanding why Kant completely rewrote the "Transcendental Deduction of the Categories" for the second edition. It is the revised definition of judgment in §19 that, in her opinion, is the outstanding feature of the new version of the transcendental deduction of the categories (Longuenesse, 1998, p. 8). Thus, according to Béatrice Longuenesse, the transcendental deduction of the categories is linked to the metaphysical deduction in a much more direct way than was the case in the first edition of 1781 (Longuenesse, 1998, p. 9).

If we look back from the transcendental deduction of the categories to the metaphysical deduction of the categories, then a sentence from §10 strike us as important:

> The same function that gives unity to the different representations **in a judgement** also gives unity to the mere synthesis of different representations **in an intuition**, which, expressed generally, is called the pure concept of understanding (KrV B104f.).

From there it also becomes understandable why the proof that is to be presented in the transcendental deduction of the categories is divided into two proof steps. While the first step of the proof concerns the logical forms of the judgment, the second step of the proof considers the subordination of the intuitions under the logical forms considered in the first step of the proof (Longuenesse, 1998, p. 9).

Bibliographic References

HENRICH, D. 'The Proof-Structure of Kant's Transcendental Deduction'. In: *The Review of Metaphysics*, vol. 22, n° 4, 1969, pp. 640-659.

HENRICH, D. *Identität und Objektivität. Eine Untersuchung über Kants transzendentale Deduktion*. Heidelberg: C. Winter, 1976.

HENRICH, D. 'Kant's Notion of a Deduction and the Methodological Background of the First Critique'. In: ECKART, Förster (ed.). *Kant's Transcendental Deductions. The three 'Critiques' and the 'Opus postumum'*. Stanford: Stanford University Press, 1989.

KANT, I. *Critique of Pure Reason*. GUYER, Paul and WOOD, Allen W. (Trans. and Eds.), The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 1998.

KOCH, F.A. *Subjektivität in Raum und Zeit*. Frankfurt am Main: V. Klostermann, 1990.

LAYWINE, A. *Kant's Transcendental Deduction. A Cosmology of Experience*. Oxford: Oxford University Press, 2020.

LONGUENESSE, B. *Kant and the Capacity to Judge*. Princeton: Princeton University Press, 1998.

PROOPS, I. 'Kant's Legal Metaphor and the Nature of a Deduction'. In: *Journal of the History of Philosophy*, vol. 41, n° 2, 2003, pp. 209-229.

STRAWSON, P.F. *The Bounds of Sense. An Essay on Kant's Critique of Pure Reason*. London, New

York: Routledge, 1989 (first published 1966).

STRAWSON, P.F. *Individuals. An Essay in descriptive Metaphysics*. London, New York: Routledge, 1971.

THÖLE, B. *Kant und das Problem der Gesetzmäßigkeit der Natur*. Berlin, New York: W. de Gruyter, 1991.

TUSCHLING, B. (ed.) *Probleme der "Kritik der reinen Vernunft"*. Kant-Tagung Marburg 1981. Berlin, New York: W. de Gruyter, 1984.

Recebido 11/2022 | Aprovado 02/2023

Enlightenment as the normative principle of social rationality

[O Esclarecimento como princípio normativo da racionalidade social]

Joel T. Klein[1]

Universidade Federal do Paraná (Curitiba, Brasil) / CNPq

DOI: 10.5380/sk.v21i1.91982

Abstract

This paper argues that Kant's philosophy entails a normative concept of social rationality distinct from other normative concepts such as moral and juridical rationality. I also argue that the normative demands of social rationality are embodied in the regulative idea of enlightenment which implies a profound intersubjective and moral character that unfolds in multiple layers and perspectives.

Keywords: Social rationality; normativity; enlightenment; Kant.

1 Professor at the Philosophy Department of the Federal University of Paraná. CNPq. E-mail: joel.klein@ufpr.br This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. Project CAPES/DAAD/PROBRAL (88887.627936/2021-00): "Perspectivas Kantianas sobre as causas da irracionalidade social". I am also grateful to Jean-Christophe Merle, Cristina Consani, Alessandro Pinzani, Maria de Lourdes Borges, Frank Rettweiller, Vinícius de Figueiredo, Marcus Willaschek, and Achim Vesper, for questions and suggestions.

This article argues that the concepts of social rationality and enlightenment are closely related. The analysis and reconstruction of arguments take place in two sections. Firstly, I reconstruct Kant's concept of social rationality and showcase the specificity of its normative field and present the sources of social irrationality. In the second section, I lay out the various aspects and layers of Kant's concept of enlightenment and argue that its requirements are actually the constitutive normative principles of social rationality. I finally draw some conclusions and sketch an outline of proposal for a system of normative contexts.

1. The Kantian concept of social rationality

Does a concept of social rationality exist in Kant's philosophy? Could there be, in a Kantian perspective, a specific normative context of the use of reason that could be called social, and which would be distinct from the ethical, the juridical, the political, the epistemological, or the religious ones? In other words, does a normative principle of social rationality exist that is capable of distinguishing sociology (as a descriptive science of society) from social philosophy (a discipline that evaluates social institutions)? Can a normative context of reason exist that might be understood as essentially or mainly social?

An almost sterile search for the word social or society in Kant's writings would lead us to think that this notion was of little importance to the philosopher.[2] Perhaps the lack of a clear social normative context was behind several criticisms of epistemological and moral solipsism. However, I argue that this apparent lack of a clear social normative context is not the result of any contempt for the social but, rather, because it is omnipresent in Kant's philosophy and constitutes the internal link between the other normative contexts and principles.

The following passage is an important starting point for analyzing the notion of social:

> *In the human being* (as the only rational creature on earth), *those predispositions whose goal is the use of his reason were to develop completely only in the species, but not in the individual*. Reason in a creature is a faculty of extending the rules and aims of the use of all its powers far beyond natural instinct, and it knows no boundaries to its project. But reason itself does not operate instinctively, but rather needs attempts, practice and instruction in order gradually to progress from one stage of insight to another. Hence every human being would have to live exceedingly long in order to learn how he is to make a complete use of all his natural predispositions; or if nature has only set the term of his life as short (as has actually happened), then nature perhaps needs an immense series of generations, each of which transmits its enlightenment [*ihre Aufklärung*] to the next, in order to finally to propel its germs in our species to that stage of development which is completely suited to its aim. And this point in time must be, at least in the idea of human being, the goal of his endeavors, because otherwise the natural predispositions would have to be regarded for the most part as in vain and purposeless (Kant, IaG, AA 08:19).

The essay *Idea for a universal history with a cosmopolitan aim* is framed by a teleological and practical approach, which structures different uses of reason and different dispositions. This teleology leads inevitably to the social perspective since humans are neither able to develop nor fulfill predispositions on their own. Only by living in society can the human species achieve this goal. It can be stated that reason and rationality are by their nature always social

2 Kant's Doctrine of Right can be said to express his theory of society and social rationality. In a certain sense this is true because the entire doctrine of right, especially that of private right, deals with the juridical relations that can exist even without a state. As Kant argues: "For in the state of nature, too, there can be societies compatible with rights (e.g., conjugal, paternal, domestic societies in general, as well as many others)" (MS, 06:306). However, even though all rightful relations are also social, not all social relations are subject to regulation by right. So, the relations based on right are different from social ones and are regulated by a specific normative principle. This is a way for maintaining the peculiarities and differences between legal, political, and social philosophy. Even though they are closely related, each has its specificity. I am grateful to Achim Vesper for raising this issue.

and intersubjective. If Kant fails to emphasize the concept of the social, it is not because it is irrelevant, but because it is fundamental and ubiquitous in his philosophy. The same is true for the concept of representation, which is also used throughout his work without any clear elaboration. Some might see this as a flaw to be corrected, but Kant seems to understand it as a premise justified and understood due to its explanatory and problem-solving capacity, which is a result of his anti-foundationalist approach.[3]

The development of rational dispositions is not based on instinct, but on a fundamental intersubjective process that entails discovering, learning, refining, and transmitting them through teaching. It relies on numerous attempts and is easily subject to different forms of failure. The main nature of the social lies precisely in this aspect of the development of rational predispositions, in the *ability to obtain and transmit some degree of enlightenment to future generations*. This process is closely related to the development of arts and sciences as Kant states in *Anthropology*:

> The sum total of pragmatic anthropology, in respect to the vocation of the human being and the characteristic of his formation, is the following. The human being is destined by his reason to live in a society with human beings and in it to cultivate himself, to civilize himself, and to moralize himself by means of the arts and sciences (Kant, Anth. AA 07: 324).

From these quotations, at least three characteristics of the social as related to enlightenment emerge as central. Firstly, it is related to the activity of people in a society. Secondly, the social element is rooted in a process of discovery (attempts and failures), learning (depuration, intelligibility, and retention of a correct use of a principle to produce a cognition, an action or a skill), and transmission (teaching and learning) of enlightenment. Finally, its aim must be the qualitative development of the different uses of reason and their reciprocal relations, which are promoted by the development of science and art.

In this process, failures can occur at different moments. They can arise with respect to specific contexts, but also in the reciprocal relationship of contexts, so that social irrationality can manifest itself in myriad ways, but it *is always characterized as a failure in the process of the gradual development of a natural predisposition to the use of reason*. Assuming that this reconstruction is correct, then it may be considered that every normative context can be affected by social irrationality. However, although all forms of rationality and irrationality in the different normative contexts carry in themselves a social element, what specifically characterizes social rationality is the processual development of rational predispositions in society throughout several generations. Thus, for example, the correct use of an epistemological principle such as efficient or teleological causality is more properly characterized as having normative validity in the field of epistemology. Similarly, the proper use of the categorical imperative is a question of morality or moral normativity. It is obvious that these uses always contain a social aspect as well. However, the proper social aspect is related to the conditions for the gradual and constant development and transmission of epistemology and science, or the proper gradual development of morality in society, i.e., *for those principles to have an increasingly correct and proper use in history.*

If we use the teleological perspective to frame this view, we can say that the full development of human natural dispositions depends on the normative principle of social rationality. This principle requires a systematic manner of organizing the different predispositions or normative contexts, not only to prevent them from conflicting with each other but also so that they might enhance each other. This systematic organization must be seen as analogous to organic systems, while, on the other hand, from the perspective of the systematic unity of reason under the assumption of the primacy of practical reason (see Kant, *KpV* AA 05:119ff.). Thus, the full development of natural predispositions regarding reason *is systemically and fundamentally moral*. The systematic and moral aspect is understood by the manner in which the sciences and the arts

3 Here I follow O'Neill's (2015) constructivist interpretation of Kant.

develop, while the economic structure of society should also be in accordance with the moral requirements of reason.

After establishing the nature of the end, "the social" can be claimed to address the adequate means for the realization of that end, in other words, the process of the appropriate development of rational predispositions. The quest for the correct thing to do in order to achieve something is an aspect of practical philosophy. Rephrasing, the "how to do something" directed towards the means is a proper issue of hypothetical imperatives, both those of instrumental rationality (how to use things to achieve a certain end) and of prudential rationality (how to use other human beings to achieve a certain end), under a specific moral point of view.[4] In other words, the proper normative feature of social rationality has to do with the appropriate means for the full development of natural predispositions of reason in the human species. *It is a matter of both instrumental and prudential rationality*, which is subject to that specific moral end.

The normativity of social rationality is intrinsically related to the instrumental and prudential perspectives, and yet, no direct equivalence can be drawn between them. Some instrumental and prudential reasonings are *too narrow in scope* and can *contradict* the end of nature, which has a broad demand for systematicity and morality. Such an overly narrow instrumental and prudential perspective would actually produce an incomplete and partial viewpoint that is incompatible with the very idea of reason as a system of normative contexts subject to morality. These partial views would instead be an expression of social irrationality. So, *the criterion of social rationality is only normatively legitimate under the conditions of systematicity and moral unity*. This means that scientific and artistic development should not collide with the demands of morality but must always promote it as fully as possible.

Once the criteria of instrumental and prudential reasoning are encompassed within the higher requirement of *systematicity and moral unity*, then a new location appears for a critical concept of happiness. The human being "*should produce everything that goes beyond the mechanical arrangement of his animal existence entirely out of himself, and participate in no other happiness or perfection than that which he has procured for himself free from instinct through his own reason*" (Kant, IaG, AA 08: 20). Human happiness is not a natural one, because it cannot be achieved by instincts nor given by nature. The human being "have [to] labored himself from the greatest crudity to the height of the greatest skillfulness, the inner perfection of his mode of thought, and (as far as is possible on earth) thereby to happiness" (Kant, IaG, AA 08: 20). Thus, the critical concept of happiness is encompassed within the conditions of merit and self-esteem.

Social rationality is, then, how enlightenment is acquired, transmitted, and developed over generations, during which the human being "should labor and work himself up so far that he might make himself worthy of well-being through his conduct of life" (Kant, IaG, AA 08: 20). By way of analogy, social rationality can be said to be a kind of engineering dedicated to finding a self-conscious way to "bring up higher the edifice which was nature's aim". This process generally takes place by way of blind groping. Social rationality requires that we discover the optimal way for this process to happen. If a direct path cannot be found to promote it, the obstacles can at least be identified, and attempts made to mitigate them.

1.1. Sources of social irrationality

Before continuing our search for the sources of social irrationality, it is important to distinguishing failures and mistakes from irrationality. The former result from the improper use of some principle, while the latter is a *manner of thinking and acting* grounded on an erroneous principle.[5] So, rationality and irrationality have to do with methodological concerns about acting on inappropriate or false principles.

4 Regarding the differences and the relationships between the moral and the prudential see Klein (2021).

5 Kant calls attention to this when portraying types of egoism. This issue is further dealt with below.

Sources of irrationality can be located in the tension between animality and human rationality, or between the tendencies toward impulses and physical subjective conditions and the rational predispositions, as well as in the crude manner in which humans find a way to balance those different tendencies and predispositions in a systematic and moral unity. As Kant says in Anthropology,

> No matter how great his animal tendency may be to give himself over passively to the impulses of ease and good living, which he calls happiness, he is still destined to make himself worthy of humanity by actively struggling with the obstacles that cling to him because of the crudity of his nature (Kant, Anth. AA 07: 325).

When not critically enlightened and disciplined, reason becomes the servant of these impulses and aims merely to satisfy them. It builds for itself principles that cannot be normatively valid since they have not been critically justified. This is how different types of egoism (logical, aesthetic, and moral, see Kant, Anth. AA 07: 128ff) or even the passions arise (see Kant, Anth. AA 07: 265ff.). In other words, certain impulses of our animality and sensibility tend to guide reason, which then becomes the servant of the passions by assuming and creating for itself strange and illegitimate principles. Social irrationality is this servitude and partiality of reason to animalistic nature and tendencies.

In *Conjectural beginning of human history*, Kant provides three examples of the "conflict between the striving of humanity toward its moral vocation, on the one side, and the unalterable following of the laws placed in its nature for the crude and animal condition, on the other side". The first conflict refers to the difference between physical majority (sexual capacity to generate offspring) and sexual drive, on the one hand, and social majority or the capacity of an individual to support himself and his kind in civilized society on the other. A gap of at least ten years extends between physical and social maturity. However, "nature has certainly not placed instincts and faculties in living creatures so that they might struggle with and suppress them. Thus, its predisposition was not at all cut out for the moral condition, but merely for the preservation of the human species as an animal species". This conflict might only be removed in a "perfect civil constitution (the uttermost goal of nature)". Meanwhile the human species has to deal with "vices and their consequences, the manifold of human misery" (Kant, MAM, AA 08: 116f.n.).

The second conflict arises regarding the constitution of human physical force as an animal and the development of the rational faculties, namely, between "humanity as an animal species and humanity as a moral species". On this teleological framework, Kant argues that

> nature obviously has taken its decision about the life span of the human being from a viewpoint other than that of the furthering the sciences. For if the most fortunate mind stands at the brick of the great discoveries he might hope for from his skill and experience, his age steps in; he becomes dull and must leave it to a second generation (which begins again from ABC and has to traverse again the whole stretch that had already been gone through) to add a next span to the progress of culture (Kant, MAM, AA 08: 117n.).

A third conflict is also rooted in the rational demand for moral and epistemological equality, on the one hand, and the civil inequality produced by the contingent and violent manner in which human societies have established their economic and material relations. In this sense, "the human being was to labor himself out of the crudity of his natural predispositions by himself, and yet was to take care not to offend against them even as he elevates himself above them – a skill that he can expect to acquire only late and after many misbegotten attempts". Meanwhile, humanity suffers the "ills that it inflicts on itself from its lack of experience" (Kant, MAM, AA 08: 118n.).

A dialectic exists between the animal and the rational parts of the human being. Natural animalistic tendencies should not be considered wrong or devoid of purpose. They can be understood, instead, as a natural force that pushes individuals and the human species toward

survival and dispersion. However, inexperience in finding the balance between these animal tendencies and the uses and requirements of reason gives rise to countless miseries and vices that hinder the development of both rational dispositions and the adequate satisfaction of animal and sensitive inclinations.

This dialectic of principles, which do not necessarily oppose each other but which clash in certain situations, is also seen teleologically as a way to promote the culture of reason itself. It is in this context that the thesis of unsociable sociability comes into play.[6] The most elaborate formulation of this position appears in the fourth proposition of the *Idea* and is understood as an anthropological statement about human nature. The unsociable tendencies serve to drive the development of rational dispositions. However, since this process lacks an appropriate culture of discipline, said dispositions end up leading to situations and behaviors that compromise the proper uses of rationality itself.

When the development of the various rational faculties is guided by animal or unsociable inclinations, reason not only becomes a slave, but it also becomes evil, since it submits to principles that lack legitimacy and which therefore tend to corrupt and frustrate its development. The sources of social irrationality can be said to be identifiable (through reflection) and even subject to neutralization (through correct discipline), but its roots cannot be eliminated.

This dialectic also affects philosophy, and in order to neutralize its negative effects, the philosophical investigation has to clearly determine the different types of validity and legitimate contexts for the use of the principles of reason. This is precisely Kant's aim in what he calls critical philosophy.

2. Enlightenment as the normative principle of social rationality

A practical teleological point of view, that tries to construct a moral and rational meaning for the existence of a rational species on earth[7], have to consider how each generation "transmits its enlightenment to the next, in order to finally to propel its germs in our species to that stage of development which is completely suited to its aim" (Kant, *Ia*G, AA 08:19). By taking the opposite path namely, the assumption that this practical project has no value or meaning, one "would remove all practical principles and thereby bring nature, whose wisdom in the judgment of all remaining arrangements must otherwise serve as a principle, under the suspicion that in the case of the human being alone it is a childish play" (Kant, *Ia*G, AA 08:19). So, the meaning and normative contented of what is enlightenment is at the center of what has to be consider social rationality. But what is enlightenment?

Kant defines enlightenment as

> the human being's emergence from his self-incurred minority. Minority is inability to make use of one's own understanding without direction from another. This minority is self-incurred when its cause lies not in lack of understanding but in lack of resolution and courage to use it without direction from another. *Sapere aude*! Have courage to make use of your own understanding! is thus the motto of enlightenment (Kant, WA, AA 08: 35).

This may be one of the most famous but also one of the most underestimated and superficially interpreted passages of the Kantian *corpus*. The essay *Enlightenment* has several layers whose meanings only come to light against the background of Kant's major writings and his idea of philosophy. I draw several distinctions below both raised and assumed throughout this essay so

6 Regarding the concept of sociable unsociability, see Klein (2013).

7 On the Kantian project of constructing an expanded moral image of the world from a practical reflexive teleology, see Klein 2016 and 2019.

that its full complexity may be understood and evaluated.[8]

2.1 The relation of two contexts: subjective and intersubjective

The concept of enlightenment is elaborated in two compatible but distinct contexts. Subjectively speaking, it relates to a particular agent's decision and behavior, while the intersubjective perspective deals with conditions, historical and societal, that make up social life. The subjective and intersubjective contexts are in constant interaction.

Regarding the intersubjective point of view, Kant argues against the idea that minority results from a natural and immutable condition inherent to the human being, which is supposedly caused by a malevolent nature or by a revengeful God. If neither nature nor providence can be blamed for the human situation, then the responsibility must inevitably fall to human beings themselves. In other words, human beings must be seen to be historically responsible for their condition. In this sense, wars, hunger, misery, ignorance, and most human sufferings result from human actions and decisions, which could have been different. Thus, they are neither the results of natural laws nor are they immutable facts.[9]

The intersubjective aspect constitutes the social and historical context of minority, and has by far the greater impact. Webs of social relations reinforce behavior, in which the absence of certain institutions and the actions of others maintain and further minority. A group educated in a desire for glory, power, and honor grows attached to playing the role of guardians of the people. Meanwhile, out of laziness or cowardice, others enact the role of domesticated animals. Nevertheless, the mistake must be avoided of thinking that the guardians are enlightened while those that are cared for are minors. Both groups are trapped in minority, yet the ills manifest differently. Those in command are slaves of their passions, while those who obey are used as mere things. An objectification takes root on both sides underpinning ways of thinking and feeling that appear to be a second nature and that engender affection. Minority is fostered by a dialectics strengthening the relation between guardians and their domestic animals, so much so that when it achieves a certain level it can truly incapacitate self-thinking. At work here is a kind of psychology, sociology, and anthropology of domination and minority.

It would be incorrect to say, however, that laziness and cowardice affect only those who obey. These conditions are equally strong in the guardians, who become accustomed to merely repeating what they have learned. Moreover, they constantly suffer the effects of their own web of prejudices. As Kant noted

> the public, which was previously put under this yoke by the guardians, may subsequently itself compel them to remain under it, if the public is suitably stirred up by some of its guardians who are themselves incapable of any enlightenment; so harmful is it to implant prejudices, because they finally take their revenge on the very people who, or whose predecessors, were their authors (Kant, WA, AA 08: 36).

Collective responsibility walks hand in hand with that of the individual, and Kant never once indicated any hierarchy between the two. It is more likely that their importance would vary situation to situation. People born and raised in unfavorable conditions would have less

8 Some arguments developed in the next pages are found in Klein (2009). However, I have worked them out in several aspects in what follows.

9 Cf. "[The human being] is not justified to ascribe his own misdeeds to an original crime of his ancestral parents, through which a propensity to similar transgressions has supposedly become hereditary in their posterity (for voluntary actions cannot bring with them anything hereditary); but rather that he must recognize with full right what they did as having been done by himself and attribute the responsibility for all ills arising from the misuse of his reason entirely to himself, since he may very well become conscious of the fact that he would have behaved in precisely the same way under the same circumstances and would have begun the first use of reason that way (even against nature hint's). When this point about moral ills is set right, then the physical ills proper can hardly amount to a surplus in our favor in the balance of merit and guilt" (Kant, MAM, AA 08: 123). See also IaG, specifically, the third proposition.

responsibility for their minority than others in their society.[10] In this sense, to be enlightened is not at all only a matter of decision and courage, because it is also an issue of capacity, whose development is strongly related to social conditions.[11] However, from the acknowledgment of the relevance of social conditions, it does not follow that the "world of readers" should be or is restricted to a kind of economic upper class or to a kind of technocracy.[12] In this sense, Willaschek has correctly pointed out that the restriction of the freedom to the public use of reason to the scholarly world does not imply a social exclusion of "the great unthinking masses" (Kant, WA AA 08:35),

> but results from the fact that a certain level of education is required to read and understand the publications in question and thus to participate competently in public discourse. The more people who have the necessary education, however, the larger the audience, which thus tends to approach a "society of the citizens of the world" (Kant, WA AA 08:37) (Willaschek (2023, 128). Translation by the author).

If Kant's moral philosophy is taken as a parameter, it should be recalled that his intention was never to present a theoretical ethical theory for evaluating individuals' actions. His was a normative moral theory aimed at evaluating principles, in particular the maxims of agents. The same holds true in *Enlightenment*. Kant is not interested in presenting criteria for identifying who has more or less responsibility in any particular situation. He defends, instead, that minority is a human responsibility, with both its social and individual dimensions, so said responsibility is shared both by the society and by singular individuals.

Whereas the social is of greater impact both for fostering enlightenment and perpetuating minority, it is the individual who hold the potential for starting a change. The new condition always starts with a few individuals with peculiar capacities and qualities that allow them to overthrow the yoke of minority.

> For there will always be a few independent thinkers, even among the established guardians of the great masses, who, after having themselves cast off the yoke of minority, will disseminate the spirit of a rational valuing of one's own worth and of the calling of each individual to think for himself (Kant, WA, AA 08: 36).

If the process depends necessarily on the initiative of a few independent thinkers, it is the social context that will assure the success or failure of their efforts. "Social context" should be understood to be comprised of all political and state institutions. It is the state that has greater conditions to amplify and propagate its effects. In Kant's words,

> one must admit that the greatest ills that oppress civilized peoples steam from war, yet to be sure less from one that actually is or has been than from the never relenting and even ceaselessly increasing *armament* for future war. To this are applied all forces of the state, all fruits of its culture that could be used for a still greater culture (Kant, MAM, AA 08:121).

10 In the Metaphysics of Morals Kant stresses that the degree of guilt is based on the obstacles from the grounds of duty: "the possibility to talk of greater or less guilt based on the less the natural obstacles and the greater the obstacle from grounds of duty, so much the more is a transgression to be imputed (as culpable). - Hence the state of mind of the subject, whether he committed the deed in a state of agitation or with cool deliberation, makes a difference in imputation, which has results" (Kant, MS, 06:228).

11 In this sense I agree with Willaschek, who writes: Kant "is therefore by no means of the opinion that enlightenment and thinking for oneself is a mere matter of decision. The opponents of vaccination may have the courage to use their own minds without the guidance of others, but this does not yet mean that they are capable of doing so" (Willaschek, 2023, 130, translation by the author).

12 Kant's rejection of technocracy and the restriction to enlightenment to a ruling class is related to his position against Plato's idea of a Philosopher King: "That kings should philosophize or philosophers become kings is not to be expected, but it is also not to be wished for, since possession of power unavoidably corrupts the free judgment of reason. But that kings or royal peoples (ruling themselves by laws of equality) should not let the class of philosophers disappear or be silent but should let it speak publicly is indispensable to both, so that light may be thrown on their business; and, because this class is by its nature incapable of forming seditious factions or clubs, it cannot be suspected oi spreading propaganda" (Kant, ZeF, AA 08: 369).

> As long, however, as states apply all their powers to their vain and violent aims of expansion and thus ceaselessly constrain the slow endeavor of the inner formation of their citizens' mode of thought, also withdrawing with this aim all support from it, nothing of this kind is to be expected [the moralization of humanity], because it will require a long inner labor of every commonwealth for the education of its citizens (Kant, IaG, AA 08: 26.).
>
> For while the people feel that the costs for education of their youth ought to be borne, not by them, but by the state, the state for its part has no money left (as Büsching complains) for the salaries of its teachers who are capable and zealously devoted to their spheres of duty, since it uses all the money for war. Rather, the whole mechanism of this education has no coherence if it is not designed in agreement with a well-weighed plan of the sovereign power, put into play according to the purpose of this plan, and steadily maintained therein; to this end it might well behoove the state likewise to reform itself from time to time and, attempting evolution instead of revolution, progress perpetually toward the better (Kant, SF, AA 07:93).

The social and intersubjective aspects can be said to take precedence in the sense that they are decisive for the comprehensive and everlasting effects of a way of thinking. In this respect, Kant's evaluation is clear: If throughout history the human species had invested all its effort and money in education instead of war, it would have already made the transition from an age of enlightenment to an enlightened age. Nevertheless, it is the subjective and personal point of view that has historical and epistemological precedence. Understanding this order of relation is key because, although the social perspective is absolutely central to enlightenment, it lacks capacity to take responsibility for all the work. At the end of the day, each of us has to free ourselves, and the most that society can do is to foster and increase our possibilities. Only in this sense should Kant's philosophy be understood to be subject-centered. This differs greatly from the Hegelian interpretation that claims that Kant's philosophy fails to recognize the importance of the social and intersubjective sphere or that it bears a monist and egoist conception of reason and rationality.

2.2 Precepts and formulas versus principles

The first definition of enlightenment is having the courage and the ability to make use of one's own understanding. But what does this mean? The notion of thinking for oneself does not elucidate much, because it seems that we are walking in circles. Yet, in the third paragraph, an important hint appears, namely that "precepts and formulas, those mechanical instruments of a rational use, or rather misuse, of his natural endowments, are the ball and chain of an everlasting minority" (Kant, WA, AA 08: 36). What does it mean to be chained to precepts and formulas as mechanical instruments that do not allow "free movement"? Could Kant be claiming that thinking freely might allow one to give up thinking according to rules? This is not the case, because, as argued in section 1 above, they are the two sides of the same coin.

Making free use of reason and understanding does not mean thinking without rules. On the contrary, it means engaging in reason by following merely the rules given by reason, despite and independently of precepts and formulas laid down by religions, cultures, and customs. In *What does it mean to orient oneself in thinking?*, Kant claims that "freedom in thinking signifies the subjection of reason to no laws except *those which it gives itself*; and its opposite is the maxim of a *lawless use* of reason (in order, as genius supposes, to see further than one can under the limitation of laws)" (Kant, WDO, AA 08:145). So, for Kant, freedom means following the laws that reason gives itself, while the maxim of lawless thinking will end in a position that Kant calls libertinism, i.e., "the principle of recognizing no duty at all" (Kant, WDO, AA 08:146).

This difference also arises in Kant's debate with his former student Herder. Kant begins his first review of Herder's *Ideas for the philosophy of the history of humanity* by praising his sagacity and fruitful mind, but he then goes on to harshly criticize his student's methodology. His way of analyzing and presenting ideas, Kant argues, makes them "less capable of communication",

according to which Kant means "a logical precision in the determination of concepts or a careful distinction and proof of principles". Rather than follow this procedure, Herder choses the path of a

> sagacity adept in the discovery of analogies and a power of imagination bold in the use of them, combined with the skill in captivating its object, always held at an obscure distance, through feelings and sensations, which allow more to be surmised about them as the effects of a great content of thoughts, or as hints pregnant with meaning, than cold judgment would ever encounter in them outright (Kant, *RezHerder,* AA 08: 45).

So, we can fairly establish a distinction between precepts and formulas, on the one hand, and the autonomous principles and rational rules, on the other. Giving up these self-imposed rules would imply an inability to communicate thoughts, while precepts and formulas have no authority based on our self-reflection but come from external sources. Kant writes, "it is so comfortable to be a minor! If I have a book that understands for me, a spiritual advisor who has a conscience for me, a doctor who decides upon a regimen for me, and so forth, I need not trouble myself at all. I need not think, if only I can pay; others will readily undertake the irksome business for me." (Kant, WA, AA 08: 35) The submission to foreign *supervision* does not mean taking others' opinions and arguments into account. Rejecting arguments of authority is not the same as rejecting the authority and coherence of others' thoughts. This is expressed in the moto: "*Caesar non est super grammaticos*" (Kant, WA, AA 08: 40). We must draw a distinction then between *authoritative thinking and authoritarian thinking*. In other words, although striving for enlightenment implies not accepting arguments of authority, it also implies a rejection of logical egoism, which can only lead to relativism and skepticism. In Kant's lectures of logic, we find the following definitions of logical egoism:

> This so-called logical egoism consists, then, in nothing but the presumed but often false self-sufficiency of our understanding, existing for itself and, so to say, isolated, where one believes he knows enough by himself, and believes he is infallibly correct and incorrigible in all his judgments. And we easily see that this conceited mode of thought is not only completely ridiculous but is even most contrary to real humanity (Kant, V- Lo/Blomberg, AA 24:151).

> *egoismus logicus*[,] seems initially actually to be allowed. One has no need to consult other judgments, of course, when one cognizes something correctly oneself;] one need not seek to bring about an agreement of others with the cognition that one has oneself. But on closer investigation we find that one cannot be certain whether one has judged rightly or not if one has not compared his judgments with the judgments of others and tested them on the understanding of others. For a cognition is not correct when it agrees with my private understanding but when it agrees with the universal laws of the understanding of all men. (...) Logical egoism is opposed, however, to another prejudice, namely, where one builds all his judgments on the reason of others, judges nothing himself, but merely imitates, in that one elevates others above oneself and trusts nothing to oneself. To speak briefly, these are the prejudices of the lazy (Kant, V- Lo/Blomberg, AA 24:187f., see also Log, AA 09:80).

> Of logical egoism {Logical egoism is a selfish prejudice} This is not merely conceit but rather a kind of logical principle, which takes as dispensable the criterion of truth, to compare one's opinions with those of other men (Kant, V-Lo/Dohna, AA 24:740).

According to Kant, then, the minority of laziness cannot be surpassed by logical egoism, because egoism falls short in respecting "the universal laws of the understanding of all men". To a certain extent, we can say that enlightenment opposes both the laziness of accepting everything from others and the arrogance of the logical egoist. Agreeing with the laws of understanding and reason is not something that can be decided by statistics, such as conducting a survey asking how many people believe or do not believe that the Earth is flat or round. Instead, this question is something that has to be proven through argument.

Being enlightened has less to do with the content of the thinking or believing than with

the manner it is done. The issue is not so much the content of the book, the specific moral commandments, or advice of the priest or even about what specific diet is to be followed. It is, above all, about the *way* or the *how* one assumes these contents. Enlightenment is not about replacing some prejudices with others. For this reason, Kant is skeptical of revolutions. They "may well bring about a falling off of personal despotism and of avaricious or tyrannical oppression, but never a true reform in one's way of thinking; instead new prejudices will serve just as well as old ones to harness the great unthinking masses" (Kant, WA, AA 08: 36). Enlightenment is about taking a critical position about one's own assumptions. Before taking something-to-be-true, it must be subjected to critical evaluation. In this sense, being enlightened is not about knowing something, it is, rather, about being capable of publicly justifying it, which can only happen if one has seriously reflected on it.

> Thinking for oneself means seeking the supreme touchstone of truth in oneself (i.e., in one's own reason); and the maxim of always thinking for oneself is *enlightenment*. Now there is less to this than people imagine when they place enlightenment in the acquisition of information; for it is rather a negative principle in the use of one's faculty of cognition, and often he who is richest in information is the least enlightened in the use he makes of it. To make use of one's own reason means no more than to ask oneself, whenever one is supposed to assume something, whether one could find it feasible to make the ground or the rule on which one assumes it into a universal principle for the use of reason. This test is one that everyone can apply to himself; and with this examination he will see superstition and enthusiasm disappear, even if he falls far short of having the information to refute them on objective grounds. For he is using merely the maxim of reason's self-preservation. Thus it is quite easy to ground enlightenment in individual subjects through their education; one must only begin early to accustom young minds to this reflection. But to enlighten an age is very slow and arduous; for there are external obstacles which in part forbid this manner of education and in part make it more difficult (Kant, WDO, AA 08: 146n.).

> For the sake of practice in thinking for ourselves, or philosophizing, we will have to look more to the method for the use of our understanding than to the propositions themselves at which we have arrived through this method (Kant, Log, AA 09: 26).

But how can precepts or formulas be distinguished from self-imposed principles of reason in a particular situation? To a certain extent, this question can be easily answered. General logic is a necessary, though insufficient, condition for thinking correctly. As was pointed out in the first section, logic is not a descriptive, but a normative discipline which uses understanding and reason to bring them into alignment.

Logic is, however, only the formal criteria for evaluating our thinking. Philosophy (in a broader sense) provides more substantive criteria. In *Critique of pure reason*, philosophy as a *conceptus cosmicus* is defined as the "science of the relation of all cognition to the essential ends of human reason (*teleologia rationis humanae*), and the philosopher is not an artist of reason but the legislator of human reason." (Kant, KrV, B 867) In this case, the philosopher represents a normative ideal, something that never will be fully empirically realized. This does not mean that philosophers have to be considered as oracles, on the contrary. The key here is the role that philosophy has to perform, namely, to investigate the rules and principles that enable a rightful way of thinking and reasoning. In order to fulfill its task, philosophy must be granted the freedom of the public use of reason. "So, the philosophy faculty, because it must answer for the truth of the teachings it is to adopt or even allow, must be conceived as free and subject only to laws given by reason, not by the government" (Kant, *SF*, AA 07: 27). Therefore, enlightenment is not the refusal of any kind of authority, because the grammarians still hold legitimacy based on their knowledge and specialties. Their legitimacy supersedes even *Caesar's authority*. Their work is subject to challenge, but neither from outside the field nor by using procedures and methods alien to the subject at hand.

2.3 The difference between the private and the public use of reason and its claims of freedom

> The *public* use of one's reason must always be free, and it alone can bring about enlightenment among human beings; the *private* use of one's reason may, however, often be very narrowly restricted without this particularly hindering the progress of enlightenment. But by the public use of one's own reason I understand that use which someone makes of it *as a scholar* before the entire public of the *world of readers*. What I call the private use of reason is that which one may make of it in a certain civil post or office with which he is entrusted (Kant, WA, AA 08: 37).

It is important to note that the public use of reason implies two requirements. Firstly, the debate among a free community of equals[13] as a world of readers must always proceed by way of rational argumentation. Secondly it must be carried out according to certain procedures. When Kant claims that public use is only possible *as long as one behaves as a scholar*, the point here is not that only scholars can make a public use of reason, as this is often understood. Rather than proposing a technocracy, or a meritocracy based on erudition, Kant suggests appealing to the ideal of a scholar, in which the points and positions must be presented with solid arguments instead of merely irreflective and random opinions.

The world of readers and scholars' behavior establish necessary criteria for fostering enlightenment, as they imply a commitment not only to a manner of argument but also to the whole (or most of such) of the accumulated knowledge in that field. These conditions aim to avoid what was characterized in the *Conflict of Faculties* as an illegal conflict (Cf. Kant, *SF,* AA 07: 29-32). Said illegality arises from the illegitimate appeal to arguments based on authority (tradition or religion) or the prejudices and sentiments of the masses that are ignorant of the issue at hand. As a result, the conflict is relegated to the status of a mere dispute or discussion whose goal is only to win or impose a position.

The freedom for a public use of reason is absolutely necessary for enlightenment, and no civil compulsion in this matter is morally legitimate, unless intended to enable the freedom of the public use of reason itself. So, for Kant,

> the freedom to think is opposed first of all to civil compulsion. Of course, it is said that the freedom to speak or to write could be taken from us by a superior power, but the freedom to think cannot be. Yet how much and how correctly would we think if we did not think as it were in community with others to whom we communicate our thoughts, and who communicate theirs with us! Thus, one can very well say that this external power which wrenches away people's freedom publicly to communicate their thoughts also takes from them the freedom to think - that single gem remaining to us in the midst of all the burdens of civil life, through which alone we can devise means of overcoming all the evils of our condition (Kant, WDO, AA 08:144).

Here again it becomes evident that the freedom of the public use is fundamental, because only with it can we communicate our thoughts and share methods for testing their legitimacy. Returning to an issue raised above, the enlightenment can be defined as being *a way* of possessing knowledge. Enlightenment cannot exist without objective knowledge, which is not reached simply by external agreement (as in a statistical survey of those supporting or opposing an issue) but is an objective and intersubjective construction that must follow certain principles, which can only become clear and be kept in check when someone behaves as a scholar or a philosopher in the community of free equal citizens of the world of readers.

However, Willaschek is right in stressing that Kant's statement that "a public should enlighten itself is more possible; indeed this is almost inevitable, if only it is left its freedom" (Kant, WA, 08:36) carries a kind "a rhetorical exaggeration that owes much to the political ambition of Kant's essay", but which is "relativized by Kant himself" (Willaschek 2023, 130.

13 Cf. *KrV* B 766. See Höffe, 1996 and 2007.

Translation by the author). That the freedom alone is not sufficient is made clear by Kant himself when he stresses the importance of public education and the pernicious role of war. This rhetorical point should be seen as a strategy or as an example of Kant's political engagement in trying to convince the political elite that freedom of thought will not bring social chaos and disorder. So, we could summarize Kant's defense on the freedom of the public use of reason as covering a full spectrum of strategies and arguments:

a) It is an inalienable right of humanity (Kant, WA, 08: 39; TP 08: 304; MM 06: 327).

b) It is a necessary, although not sufficient condition for the enlightenment. There are also other important social, such as public education (see Kant, MAM, AA 08:121; IaG, AA 08: 26; SF, 07:93).

c) It is "the sole palladium of the people's rights" (Kant, TP AA 08: 304).

d) It is based in "a natural calling of humanity to communicate with one another, especially in what concerns people generally" (Kant, TP AA 08: 305).

e) It does not endanger social and political stability[14].

f) Fostering the freedom of the public use of reason would be the only secure way by which future generations will truly honor the past "head of states as well as their servants" (Kant, IaG, AA 08: 31).

g) Freedom of the public use of reason is a necessary political condition in order to assume that the "head of state" is truly interested in the commonwealth and is not in contradiction with its intention. So, denying the "freedom of the pen" means withholding from the head of state "all knowledge of matters that he himself would change if he knew about them and to put him in contradiction with himself" (Kant, TP AA 08: 304).

h) Should one assume that the head of state does not need the information and opinions that result from the freedom of the public use of reason, this would be the same "to assume that the head of state could never err or be ignorant of something" and, therefore, it "would be to represent him as favored with divine inspiration and raised above humanity" (Kant, TP AA 08: 304).

i) Those who foster suspicious on the head of state regarding the freedom of the public use of reason can only manage to awake "in him mistrust of his own power or even hatred of his people" (Kant, TP AA 08: 304).

j) The existence of "the mechanism of the state constitution to coercive laws" without the freedom of the public use of reason leads to creation of secret societies, which undermines and jeopardizes the government (Kant, TP AA08: 305).

So, from all those arguments and quotations above, is transparent that one of the main concerns that Kant had in his defense of the public use of reason was about censorship.

Alternatively, the freedom of private use of reason, must always be restricted in some measure, otherwise life in a society would be impossible. Therefore, against what is usually understood in the literature, the private use of reason does not only concern the use of reason made by public officials and functionaries in the exercise of their office or function[15]. The

14 See: "But the frame of mind of a head of state who favors the first goes still further and sees that even with respect to his legislation there is no danger in allowing his subjects to make public use of their own reason and to publish to the world their thoughts about a better way of formulating it, even with candid criticism of that already given; we have a shining example of this, in which no monarch has yet surpassed the one whom we honor" (WA, AA 08:41).

15 See: "The only exceptions to this freedom of opinion and of the press are statements made by officials and functionaries in the exercise of their office or function. This is what Kant calls, terminologically

private use of reason reaches all social functions that an individual might exercise, including his position as citizen. Is precisely the distinction between freedom of the private and public use of reason that underlies this example given by Kant right after he discusses the cases of an officer, a tax official and a clergyman, namely:

> A citizen cannot refuse to pay the taxes imposed upon him; an impertinent censure of such levies when he is to pay them may even be punished as a scandal (which could occasion general insubordination). But the same citizen does not act against the duty of a citizen when, as a scholar, he publicly expresses his thoughts about the inappropriateness or even injustice of such decrees (Kant, WA, AA 08: 37).

The agreements owing from the field of the public use of reason cannot be subject to any compulsion (internal or external to the specific field); they must result from free agreement. Such agreements can be slow to achieve, however. Meanwhile, our preexisting prejudices and social arrangements that guide our daily life cannot be suspended. Thus, each person *cannot* be allowed to decide how to react in the event of a disagreement. Take a very simple example: Let us assume a situation in which people disagree about whether a traffic light should be taken down at some crossroad. Until a public decision is issued, each individual must continue to stop at the red light. The restriction of the private use of reason has the same importance for enlightenment as does freedom in the public use. So, we should not underestimate the importance of limiting the freedom of the private use of reason. It is in this sense that we should understand Kant's appraisal of Frederic's motto "*Argue as much as you will and about what you will; only obey!*" (Kant, WA, AA 08:41). Kant's point here was not to praise the king but to make a conceptual point about the importance of social arrangements.

2.4 Rational and historical knowledge

The requirement of thinking for oneself must be understood in some context, otherwise it would become absurd in its unrestricted demands. Already in Kant's time, it was impossible to critically evaluate all knowledge and beliefs. Thus, another distinction must be introduced, namely those between rational and empirical cognition and between rational and historical manners of knowing:

> Cognitions of reason are opposed to historical cognitions. The former are cognitions from principles (*ex principiis*), the latter cognitions from data (*ex datis*). — A cognition can have arisen from reason and in spite of that be historical, however, as when a mere literator learns the products of someone else's reason his cognition of these products of reason is then merely historical, for example. One can distinguish cognitions, then,
>
> 1. according to their objective origin, i.e., according to the sources from which alone a cognition is possible. In this respect all cognitions are either *rational or empirical*;
>
> 2. according to their subjective origin, i.e., according to the way in which a cognition can be acquired by men. Considered from this latter viewpoint, cognitions are either *rational or historical*, however they may have arisen in themselves. Hence something that is subjectively only historical can be objectively a cognition of reason (Kant, Log, AA 09:22).

If we cross the categories of objective and subjective origin, three viable combinations arise: firstly, a knowledge that has a rational origin and is known by someone in a rational way (this is the case of a mathematician understanding a proof of some mathematical principle, or a moral philosopher who understands why some principle is moral); second, a knowledge that has rational origin but is known only historically (this is the case of someone that knows only how to use a mathematical principle to resolve a particular problem, or a moral person that knows how to follow some moral principle); the third category is that in which empirical knowledge is

somewhat confusingly, the "private use" of reason, while the use of reason in the publications of a private scholar is a 'public use'" (Willaschek, 2023, 128).

known historically (I know that China is real because I believe in what other people have said). The fourth category is contradictory, because empirical knowledge cannot be known through any kind of rational demonstration (there is no way to rationally prove that Berlin exists or that some mammals can fly). Empirical knowledge can only be known historically, and this means that I always have to believe in someone, myself or someone else. For Kant there is no unquestionable preference for my experience rather than that of others when empirical knowledge is at issue.

These distinctions are highly important, because they imply differences in the way we reach enlightenment. Moreover, a person may be enlightened in mathematics but can be minor about politics and ethics.

> With some rational cognitions it is harmful to know them merely historically, while with others it makes no difference. Thus the sailor knows the rules of navigation historically from his tables, for example, and that is enough for him. But if the jurist knows jurisprudence merely historically, then he is fully ruined as a genuine judge, and still more so as a legislator (Kant, *Log*, AA 09:22).

What distinguishes the case of the jurist from that of the sailor? The simple mechanical application of the law might work sometimes. However, without grasping the "spirit of the law" situations will always arise in which, without rational knowledge of the law, the consequences might be the complete opposite of what the law was meant to achieve.

Since it is impossible for anyone to be enlightened in every domain[16], then the following issue arises: in which fields should everyone be enlightened and in which area would it be sufficient for a particular group of people to be so? Kant argues that everyone should *strive for enlightenment* in *at least* in what concerns morality, because it is the basic building block of genuine human life[17]. He goes on to state that everyone should also be enlightened about their specific fields of activity. For example, a mathematician should be enlightened about mathematics, but a philosopher need not be and vice-versa. It would be highly problematic were a mathematician incapable of understanding the principles of her field, for she would be a minor where she is supposed to be a citizen. She is a minor because she would only have the historical knowledge of her field. Therefore, she would not be able to conduct herself as a scholar before her peers. In my case as a philosopher, however, it is not a problem if I have only a historical knowledge of mathematics. In this sense, we have to recognize that an enlightened age would not be the one in which everyone is enlightened about everything. Instead, it is a period when most people are enlightened about religion, ethics, and politics, while particular groups are enlightened in specific areas. Having taken this into account, we can say that, for an enlightened age to exist, it is not required that everyone should have a rational knowledge about rational cognition or historical knowledge about all empirical cognitions. The point here is more about what kind of knowledge we are dealing with in each case and context, so that we might assess the claims each kind of judgment has and how their justification should be conducted, and also which kind of public expectations should be met in specific cases.

16 Would this definition not conflate enlightenment with knowledge? I do not think that it is the case because having a rational cognition already entails that one does not only have the knowledge in a historical sense, but that understands in a rational way the principles on which that knowledge is build on. I am grateful to Marcus Willaschek for raising this question.

17 In this sense I interpret the following passage: "If, therefore, a system of a priori cognition from concepts alone is called metaphysics, a practical philosophy, which has not nature but freedom of choice for its object, will presuppose and require a metaphysics of morals, that is, it is itself a duty to have such a metaphysics, and every human being also has it within himself, though as a rule only in an obscure way; for without a priori principles how could he believe that he has a giving of universal law within himself?" (*MS*, AA 06: 216).

2.5 The primacy of practical reason

One of the more common misunderstandings about Kant's position on enlightenment is that it was supposed to be restricted only to a theoretical realm. One such misinterpretation is that by Adorno and Horkheimer in their *Dialectics of Enlightenment*. Their position may have been prompted by an overly narrow interpretation of Kant's motto: "Sapere aude! Have courage to make use of your own understanding". It is assumed that "understanding" refers merely to the technical concept as defined in the *Critique of pure reason*. This clearly wrong as a careful reading of the essay can easily reveal.[18]

La Rocca wonders how so many readers have failed to notice that the concept of wisdom lies at the center of Kantian philosophy. He states

> wisdom is for critical philosophy not just a word for the art of life, for the use of philosophy in life, but a focal concept of critical rationality as such, which is understood as a regulative ideal rather than a model already realized or even realizable (La Rocca 2004: 128).

Also, Brandt (2018, p. 147) stresses the point that for Kant the Enlightenment concerns essentially a moral and juridical issue[19]. A close relation therefore exists between philosophy and notions of wisdom, enlightenment, terminal end, and, thus, a concept of philosophy and enlightenment that constitutes an ideal projected by pure *practical* reason. This is confirmed by several other formulations of the motto.

> Philosophia (doctrina sapientiae) ist nicht eine Kunst von dem was aus dem Menschen zu machen ist sondern was er aus sich selbst machen soll [*sapere aude*] Versuche dich Deiner eigenen *Vernunft* zu Deinen wahren absoluten Zwecken zu bedienen (Kant, *OP,* AA 21: 117, emphasis added).

> Zu allem Wissen (Scientia) dessen sich der vernünftelnde Mensch zu seinem Wohlseyn bedienen kann ist das Selbsterkentnis (nosce te ipsum) ein Gebot der *Vernunft* welches Alles enthält**:** *sapere aude sey weise* (Kant, *OP*, AA 21: 134, emphasis added).

> To make use of one's own *reason* means no more than to ask oneself, whenever one

18 In this sense I agree with Willaschek (2023, 128): "Die verzweifelte Einschätzung, zu der Horkheimer und Adorno gelangen, ist angesichts der verzweifelten weltpolitischen Lage durchaus verständlich. Und dass die menschliche Beherrschung der Natur sich gegen den Menschen selbst wenden kann, wird man in Zeiten des anthropogenen Klimawandels kaum bestreiten können. Dass aber aufklärerisches Denken notwendigerweise zu Unterdrückung und Gewalt führt, scheint dagegen aus heutiger Sicht wenig plausibel. Die Verkürzung der Vernunft auf instrumentelle Rationalität ist eher eine Verkürzung des Aufklärungsgedankens durch Horkheimer und Adorno als eine angemessene Rekonstruktion aufklärerischen Denkens. Auch die Weise, wie die beiden Vertreter der kritischen Theorie sich in ihrem Werk auf den Begründer der kritischen Philosophie, also Kant, beziehen, beruht auf einem doppelten Missverhältnis: Sie kritisieren Kant für etwas, das er nicht gesagt und gemeint hat; und wo sie mit Recht eine These Kants kritisieren, da wird Kant gar nicht erwähnt." According to Willaschek, the critique that Adorno and Horkheimer could have rightly done to Kant was its optimism regarding the freedom of the press. I agree with Willaschek that "die Entwicklung der Zeitungslandschaft in vielen Ländern, dass staatliche garantierte Pressefreiheit keineswegs zu einem kritischen öffentlichen Diskurs, oder, in Kants Worten, zu einem aufgeklärten Publikum führen muss". However, I do not think that Kant's concept of the freedom of the press is a liberal one, but a republican instead. I have developed some aspects of it in Klein (2015). Other criticism to Adorno and Horkheimer regarding their interpretation of enlightenment is found in Brandt (2018, 148). According to him, Adorno and Horkheimer's thesis fall under a kind of performative contradiction.

19 I disagree, however, with Brandt when he thinks that the enlightenment has nothing to do with the sciences and the technique. Firstly, not only the Critique of pure reason, which has as a strong theoretical emphasis, but also the development of sciences has an important implication for the enlightenment, either in the sense of delimiting different spheres of knowledge, or in the sense of fighting against various kinds of superstition. Secondly, Brandt is wrong when he disqualifies the instrumental reason and also prudence for the interest of Kant's philosophy (see Klein, 2021). If, on the one hand, restricting enlightenment to the mere technical realm is a mistake, on the other hand, it is also an error to restrict it to the practical moral realm.

> is supposed to assume something, whether one could find it feasible to make the ground or the rule on which one assumes it into a universal principle for the use of reason (Kant, WDO, 08: 146n., emphasis added).

These formulations relate the ideal of enlightenment to reason, rather than merely to understanding, which, in turn, showcases its link to the higher faculty of practical principles. Enlightenment is always under the demands of morality. For a Kantian ideal of enlightenment, practical reason always takes primacy over the theoretical one.

3. Final Remarks

If this reconstruction is sound, enlightenment can be concluded to be a normative principle of social rationality. It requires the most appropriate means for fostering the full development of our rational predispositions, considered as a systematic and moral unity. In this sense, agency and institutions based on a principle undermining enlightenment can be seen to be an expression of social irrationality.

> For in the face of the omnipotence of nature, or rather its supreme first cause which is inaccessible to us, the human being is, in his turn, but a trifle. But for the sovereigns of his own species also to consider and treat him as such, whether by burdening him as an animal, regarding him as a mere tool of their designs, or exposing him in their conflicts with one another in order to have him massacred - that is no trifle, but a subversion of the final end of creation itself (Kant, SF, AA 07:89).

Neither social rationality nor the normative content of enlightenment assume the point of view of nature, but the perspective of how human beings should relate to each other. At any moment an asteroid could hit Earth and exterminate human life. This omnipotence of nature is not under our control, nor subject to oversight. However, that we do it to ourselves is normatively wrong. We can understand and fulfill our responsibility to treat human beings as rational moral agents. Thus, social rationality must eradicate the treatment of humans as mere tools or, worse, their exposure to suffering or massacre.

The difference between social rationality and moral and legal rationality is not easy to grasp, because they are interdependent and overlap. However, some relevant distinctions can be identified. In ethics, for example, some demands can be made that require the agent to develop its potentialities, as is performed by the duty of virtue. Ethics, however, cannot tackle questions of how society should promote the development of these predispositions on a broader scale. In contrast with ethics, the normativity of social rationality deals more to with public institutions, and it tackles a broader scope of human relations. Social philosophy, on the contrary, engages with the normative, large-scale demands that are institutionalized in social practices. It is moreover closely related to the prudential and instrumental manner of discovering the right means to an end.

In comparison with the legal and political institutions, on the other hand, the social rationality does not always fall under external coercion and is more pervasive. From the point of view of right, the state should guarantee access to education and opportunities for people to develop their rational predispositions. It is the responsibility of social rationality, however, to organize social institutions (such as the family, the school, the workplace, associations, the press, and the internet) in such a way that they may also promote that end. Social and legal normativity may overlap in the sense that the law could constrain certain practices that conflict with peoples' right. However, law cannot and should not deal with all issues. Otherwise, such a coercive machine would be required that it would probably undermine any space for enlightenment since we would be outsourcing all responsibility to the State. From a Kantian perspective, not only must the connections between these normative fields exist, but their differences must also

be emphasized because the philosopher dealing with them has to be clear about what is subject to criticism, under which normative criteria, and also about the scope of what can be proposed.

Bibliographic References

ADORNO, T. W.; HORKHEIMER, M. *Dialektik der Aufklärung*. Frankfurt am Main, 1969.

BRANDT, R. Was ist Aufklärung? Beobachtungen zur Kantischen Antwort. *Kant-Studien*, 109(1), 2018, p. 147–169.

HÖFFE, O. Eine republikanische Vernunft. Zur Kritik des Solipsismus-Vorwurfs. In: *Kant in der Dikussion der Moderne*. G. Schönrich and Y. Kato (ed.). Frankfurt a. M.: Suhrkamp., 1996, p. 396-407.

HÖFFE, O. Kants universaler Kosmopolitismus. *Deutsche Zeitschrift für Philosophie*. 55, 2007, p. 179-191.

KANT, I. *Cambridge Edition of the Works of Immanuel Kant*. Paul Guyer and Allen W. Wood (ed.). Cambridge: Cambridge University Press, 1996–.

KANT, I. *Gesammelte Schriften*. Ed. Königlich Preussische Akademie der Wissenschaften (and successors). Berlin: de Gruyter (and predecessors), 1900–.

KLEIN, J. A sociabilidade insociável e a antropologia kantiana. *Revista de Filosofia: Aurora* (PUCPR. Impresso), 25, p. 265-285, 2013.

KLEIN, J. Kant on religious intolerance. *Philosophica (Lisboa)*, 51, p. 25-38, 2018.

KLEIN, J. Freedom of the Press: A Kantian Approach. *Estudos Kantianos*, 3, p. 83-92, 2015.

KLEIN, J. On Serpents and Doves: The systematical relationship between prudence and morality in Kant's political philosophy. *Kant-Studien*, 112, p. 78-104, 2021.

KLEIN, J. T. The Highest Good and the Practical Regulative Knowledge in Kants Critique of Practical Reason. *Con-Textos Kantianos: International Journal of Philosophy*, 3, p. 210-230, 2016.

KLEIN, J. T. Kant's constitution of a moral image of the world. *Kriterion*, 60, p. 103-125, 2019.

KLEIN, J. T. A resposta kantiana à pergunta: que é esclarecimento? *Ethic@* (UFSC), 8, p. 211-227, 2009.

LAROCCA, C. Was Aufklärung sein wird. Zur Diskussion um die Aktualität eines Kantischen Konzepts. In: *Recht – Geschichte – Religion: Die Bedeutung Kants für die Gegenwart* (Deutsche Zeitschrift für Philosophie, Sonderband 9). Hrsg. von Herta Nagl-Docekal und Rudolf Langthaler. Berlin: Akademie Verlag, 2004.

O'NEILL, O. *Constructing authorities. Reason, Politics and Interpretation in Kant's Philosophy*. Cambridge University Press, 2015.

WILLASCHEK, M. Kant und das Projekt der Aufklärung heute. In: *Con-Textos Kantianos*, vol. 17, p.125-133, 2023.

Recebido 11/2022 | Aprovado 02/2023

Revisiting the Kantian legacy in Habermas: the philosophical project of modernity and decolonial critiques to rationality and cosmopolitanism

[Revisitando o legado kantiano em Habermas: o projeto filosófico da modernidade e as críticas decoloniais à racionalidade e ao cosmopolitismo]

Cristina Foroni Consani[1]

Universidade Federal do Paraná (Curitiba, Brasil) / CNPq[2]

DOI: 10.5380/sk.v21i1.91983

Abstract

This paper deals with Amy Allen's critique of Habermas's theory of modernity, democracy, and cosmopolitanism. I will focus on her arguments that touch on the role of rationality. Allen's critique of Habermas will be presented, especially where she argues that focusing on rationality is ethnocentric and promotes the political exclusion of subaltern groups. The extent to which Allen's critiques affect the emancipatory potential of Habermasian theory will next be assessed. It is finally argued that Allen's position leads to a denial of the distinction between social rationality and irrationality as legitimate criteria. The consequences of such a position for the political sphere will then be analyzed.

Keywords: Kantian legacy; Habermas; rationality; democracy; cosmopolitanism.

1 cristina.foroni@ufpr.br

2 This research was also supported by the International Cooperation Project CAPES/DAAD/PROBRAL (88887.627936/2021-00): "Perspectivas Kantianas sobre as causas da irracionalidade social".

In the field of decolonial criticism, Kantian philosophy and its heirs have been accused of ethnocentrism, Eurocentrism, and of fostering new forms of imperialism and colonialism. These criticisms rest mainly on the rationalist and universalist character of such theories and on the argument that by invoking universalism and rationalism Kantian theories end up invalidating other forms of knowledge and political manifestations that do not fit the established criteria of rationality. Thus, these theories claim that rationality should no longer be a legitimate criterion for the evaluation of knowledges or political manifestations because this is a criterion considered Western, imperialist, and neocolonial (cf. Allen, 2016, 2019; Mendieta, 2009; Mignolo, 2000; 2010; 2021; Walsh, 2018).[3] Before investigating the sources of social irrationality, this paper seeks to point out that the distinction between social rationality and irrationality is legitimate and relevant.

This article deals specifically with Allen's critique of Habermas's theory of modernity, law and democracy and cosmopolitanism outlined in her book *The End of Progress*, with a specific focus on her remarks regarding the role of rationality. First, Allen's criticism of Habermas' theory is presented as is her position that focusing on rationality is ethnocentric and promotes the political exclusion of subaltern groups. The extent to which Allen's critiques affect the emancipatory potential of Habermas's theory is secondly evaluated. Finally, Allen's position is argued to lead to a denial of the distinction between social rationality and irrationality as legitimate criteria. The political consequences of such a position are then analyzed.

1. Modernity and Rationality: Allen's Critique of Habermas

Allen is a signatory to the thesis that universalism has played a crucial role in connecting European culture with European imperialism over the centuries. In this sense, imperialism as a political project could not sustain itself without the idea of empire, which in turn is nurtured by a philosophical and cultural imaginary that justifies the political subjugation of native territories and populations through claims that such peoples are less advanced, cognitively inferior, and therefore could be subordinated. (cf. Allen, 2016, p. 1). According to Allen, Habermas, like other philosophers, is attached to the idea that certain aspects of European modernity and the

3 In this sense, Eduardo Mendieta argues that Kant's theory is incapable of inspiring critical cosmopolitanism (p. 248). Kantian cosmopolitan thought is based, instead, on pedagogy, anthropology and geography. In said texts, he is of the opinion that Kant's cosmopolitanism is obviously imperialist and colonialist because Kant would have assumed the privileges of his citizenship and location in the 18th century Austrian empire. Mendieta denies neither rationality nor universality per se, but he insists that these concepts need to be reformulated so that universality can be conceived of in such a way as to incorporate diversity and diverse forms of rationality. He bases his claims on those of Walter Mignolo. (cf. Mendieta, 2009, p. 251f.) Mignolo, in his turn, states that cosmopolitanism emerged in the 18th century as a secular version of Christianity, whose objective was to expand Western ideals and civilizations and reaffirm the colonial matrix of power (p.118). Cosmopolitanism, he explains, is sustained by four practices or arguments, namely: the management and control of subjectivity, the control of authority, the control of the economy, and the control of knowledge (p. 119f.). By analyzing history from the perspective of colonization, Mignolo asserts that the very foundation of knowledge is racist and patriarchal, since, although it recognizes the ontological equality of all individuals and peoples, it nevertheless affirms an epistemic inequality, which would have legitimized European domination over colonized peoples. The subjugation of indigenous peoples, for example, would have been legitimized, from the colonizer's point of view, by the argument of the inferiority of their knowledge or their minority (cf. Mignolo, 2010, p. 121f.). Mignolo also affirms the need to reformulate the concepts of rationality and universality (cf. Mignolo, 2000), as does Allen (cf. Allen, 2016, 2019). These authors' proposals for reformulating the criteria of rationality and universality will have to wait for a future paper. For now, it is only necessary to present such criticisms of cosmopolitanism in their Kantian aspects and to point to the controversies regarding this interpretation of Kant, since the decolonial authors base their criticisms on the texts of Anthropology, Pedagogy and Geography where Kant uses a methodology based on the description of social facts and history. Critics have not consistently worked with the texts of Kantian Moral Philosophy, in which the normative foundations of practical reason are laid down in terms of universalizing equal respect and consideration for all individuals and peoples. A critical review of Mignolo's reading of Kant can be found in: Santos, 2022, p. 154-175).

Enlightenment represent an advance over pre-modern, non-modern, or traditional forms of life, and crucially this idea plays an important role in grounding the normativity of his theory (cf. Allen, 2016, p. 3).

Allen considers Habermas' continued commitment to a theory of social evolution that positions "modernity" as the result of a practical-moral learning process, and also as progress, the problem that shapes eurocentrism in his work (cf. Allen, 2016, p. 38). Allen asserts that Habermas's chosen strategy for dealing with the grounding of the normativity, based on his theory of social evolution and his theory of modernity - both connected by a philosophy of history that has progress on the horizon, leaves him vulnerable to postcolonial critics who accuse him of Eurocentrism and informal imperialism (cf. Allen, 2016, p. 39).

Considering the development of Habermas' work since the mid-1970s, two types of rationality are identified, namely, *teleological* (instrumental and strategic) and *communicative rationality*. While teleological rationality is linked to the use of resources and effectiveness, to the means/ends relationship, moral-practical or communicative rationality is associated with the possibility of "identify and reconstruct universal conditions of possible understanding" (Habermas, 1976, p. 1). It asserts that there is a universal core to communicative competence, and that this core involves the differentiation of three distinct validity claims—claims to truth, to normative rightness, and to sincerity—and of three distinct world-relations—to the objective, intersubjective, and subjective worlds, respectively.

In this regard, Moral-practical development consists in part in a progressive decentration of worldviews and heightening of reflexivity

> [i]n both dimensions [that is, individual ego development and social evolution], development apparently leads to a growing decentration of interpretive systems and to an ever clearer categorical demarcation of the subjectivity of internal nature from the objectivity of external nature, as well as from the normativity of social reality and the intersubjectivity of linguistic reality (Habermas, 1976, p. 106).

In another excerpt, Habermas states that

> The rationality inherent in [communicative] practice is seen in the fact that a communicatively achieved agreement must be based *in the end* on reasons. And the rationality of those who participate in this communicative practice is determined by whether, if necessary, they could, *under suitable circumstances*, provide reasons for their expressions (Habermas, 1984, p. 17).

In volume 2 of *The Theory of Communicative Action*, Habermas also points out that:

> Our formal-pragmatic description of the general structure of speech acts has to draw on the pretheoretical knowledge of speakers who belong to a modern and—in a sense still to be explained more precisely—rationalized lifeworld (Habermas, 1987, p. 77).

Allen interprets these passages as expressing Habermas's attachment to a reconstructive methodology that has strong ties to a particular context. In her own words:

> this way of understanding Habermas's project overlooks the simple fact that the methodology of rational reconstruction proceeds by way of a systematic reconstruction of the intuitive knowledge of a very specific group of people, namely, 'competent members of modern societies,' where 'competent' means adult subjects who have learned to differentiate the three validity claims and have mastered the three-world structure of communication (Allen, 2016, p. 51).

As a consequence of the Habermasian position, she identifies a potential for disregarding other forms of knowledge or knowledges that are not produced or manifested according to such criteria of reflection and analysis of validity claims:

> since universal pragmatics itself rests on the prior assumption that modern, post traditional structures of communication and postconventional forms of identity are

> superior to premodern, traditional ones, that is, that they more fully realize the inherent telos of language as a medium of communication the theory cannot play this justificatory role and remains bound to a conception of reason that privileges a Western point of view (Allen, 2016, p. 52).

In this respect, Allen considers that claiming universality for this conception of rationality thus involves claiming universality for modern, Occidental understanding of the world, which in turn raises the question of how such a claim can be justified (cf. Allen, 2016, p. 54). To support her critique that the Habermasian theory of rationality and modernity has this informal imperialism, she draws on the distinction made by Habermas in volume 1 of *The Theory of Communicative Action* between mythical and modern worldviews. According to him,

> mythical worldviews are distinct from modern ones in two ways: first, they fail to differentiate the objective, social, and subjective worlds (and hence they also fail to differentiate claims to truth, normative validity, and sincerity); and, second, they do not identify themselves *as* worldviews, as cultural traditions, that is, they lack reflexivity about their own status (Habermas, 1984, p. 52).

To sum up, Allen remembers that Habermas's conception of rationality has four components:

> first, the three formal world concepts (objective, intersubjective, and subjective); second, the corresponding validity claims (truth, normative rightness, and sincerity); third, the concept of a rationally motivated agreement; and fourth, the concept of reaching understanding through speech (Allen, 2016, p. 57).

And she considers that "Habermas reconstructs largely converges in its shared commitment to the developmental superiority of modernity, even as it remains cognizant of modernity's downsides". She also argues that

> Habermas does not consider the possibility that (...) the belief in the developmental superiority of modernity does not itself have an evidentiary basis but rather rests on a questionable normative decisionism and ends up corroborating the relationship modernity and Eurocentrism (Allen, 2016, p. 60f.).

The same line of argument is presented when Allen analyzes other works by Habermas, such as discourse ethics, discursive theory of law and democracy, texts on religion and cosmopolitanism. She identifies the demands of rational argumentation and the project of constitutionalizing of international law, which, according to Habermas, would have a civilizing role, as elements that support Eurocentrism and the supremacy of one model of modernity over others.

In his discourse ethics, Habermas attempts to reformulate Kant's categorical imperative into a discursive procedure for moral argumentation (D). He states this reformulated categorical imperative in his original version of discourse ethics as follows: "Only those norms can claim to be valid that meet (or could meet) with the approval of all affected in their capacity *as participants in a practical discourse*" (Habermas, 1990, p. 66). From this basic principle of discourse ethics (D) combined with an account of the normative preconditions of argumentation in general, Habermas claims to be able to derive his basic moral principle, the principle of universalization (U) that states that a moral norm is valid if and only if "all affected can accept the consequences and the side effects its general observance can be anticipated to have for the satisfaction of everyone's interests" (Habermas, 1990, p. 65). According to this proposed reading, Habermas's neo-Kantian discourse ethics offers an independent justification of (D), which provides a procedure for determining the validity of norms generally, and a derivation of (U), which provides a procedure for determining the validity of specifically moral norms (cf. Allen, 2016, p. 61f.).

Allen, however, does not consider Habermas to be a Neo-Kantian in grounding the normativity of his theory. She points to a dependency or interrelationship between discourse

ethics and Habermasian modernity theory, and this understanding grounds her critique that Habermasian universalism is rooted in a contextualism (his theory of Western modernity) that leads him into a circularity. In her words:

> Moreover, even if we were to accept that the Kantian constructivist interpretation of the justification strategy for discourse ethics is a plausible reading of some of Habermas's texts, this reading does not fit easily into his overall theoretical system, and he quite clearly rejects it in his most recent discussions of this issue. This means that Habermas can't be a straightforward Kantian constructivist, and that discourse ethics cannot serve as the normative foundation for his theory of modernity because its plausibility rests, at least in part, on that theory (Allen, 2016, p. 64).
>
> In other words, Habermas assumes that contextualism at the metanormative level undermines universalism at the level of first-order, substantive norms; he assumes, that is, that metanormative contextualism entails first-order moral and political relativism. (Allen, 2016, p. 66).
>
> Habermas seems to have two choices here. He can either bite the bullet and accept that his moral project entails an internal elucidation of the modern moral point of view. Or he can argue that the theory of modernity and discourse ethics support each other in a coherentist fashion. This way of understanding his project quite clearly raises worries about circularity: the rational reconstruction of communicative competence presupposes the superiority of modernity, while the theory of modernity presupposes the superiority of a rationalized lifeworld. Of course, one could always accept that at this most basic level, the theory is circular, but argue that this circularity isn't vicious. (Allen, 2016, p. 66).

Allen does not see Habermas as a Neo-Kantian constructivist (as, for example, Keneth Baynes and Rainer Forst), but rather as a Neo-Hegelian reconstructivist. Even if Allen's interpretation of Habermas is correct, it does not seem, however, relevant to the point I want to isolate in this text. The central focus of decolonial critique, as received by Allen, is that rationalism sets certain standards in the epistemological (modern science and its criteria of verifiability or falsifiability) and political fields (the criteria for the presentation of reasons, arguments, and justifications, self-reflection and criticism) that could legitimately become exclusionary in relation to other forms of knowledge (those traditional) or political manifestations (those that do not present themselves in the public space from the presentation of reasons, arguments and justifications - the contraposition voice (*reason*) and noise (*emotion*) (see Eklundh, 2019, p. 63ff). In this sense, it seems to me that Kantian critical philosophy, whether in the theoretical or practical realm, supports the need for such criteria and, for this reason, it seems to me irrelevant here the discussion of how much Habermas is more Kantian or more Hegelian.

The next step in Allen's critique of Habermas turns to his cosmopolitan project. In analyzing his texts in this regard, but especially some passages from "The Divided West", she considers that

> Habermas is well aware of the ways in which Enlightenment ideals, particularly ideals of progress, modernization, and development, have been entangled with the so-called civilizing mission of the West. Nevertheless, as he makes clear in his discussion of Kant's philosophy of history, he believes that these ideals can be disentangled from their ideological roots. Although Habermas's own philosophy of history shares with Kant's the "heuristic aim" of lending the "idea of the cosmopolitan condition empirical probability and plausibility" (DW, 145), Habermas insists that, in taking up Kant's philosophical-historical project, we must "look beyond the prejudices associated with [Kant's] historical horizon" (DW, 145). These prejudices include an insensitivity to cultural differences, a blindness to the explosive force of national- ism, a "'humanist' conviction of the superiority of European civilization and the white race" (DW, 146), and a lack of awareness of the fact that "European international law" was "embedded in a common Christian culture" (DW, 146). (Allen, 2016, p. 69).

According to Allen, even though Habermas embraced *the multiple modernities paradigm*, he continued to associate heightened reflexivity with an irreversible moral-political learning

process (cf. Allen, 2016, p. 72). Her interpretation is based on Habermas's statements in "A Postsecular World Society? On the Philosophical Significance of Postsecular Consciousness and the Multicultural World Society", especially the following passages:

> we can indeed trace the, for now, last socially relevant push in the reflexivity of consciousness to Western modernity (Habermas, 2010, p. 2).

> In European modernity, we observe a further cognitive push in the *same* dimensions. We observe a sharpening of the consciousness of contingency and an extension of futural anticipation; egalitarian universalism becomes more pointed in law and morality; and there is progressive individualization. In any case, we still draw our normative self-understanding from this (disregarding short-winded, fashionable denials) (Habermas, 2010, p. 8).

Based on these excerpts she concludes the following:

> So, even in the face of the charge of Eurocentrism, Habermas remains committed to a progressive view of history according to which European modernity represents a moral-political advance over premodern forms of life. And even as he recognizes the need to rethink the central assumptions of modernization theory, he insists that we shouldn't throw the baby out with the bathwater by rejecting the idea of European modernity as a privileged example of sociocultural learning processes in an effort to avoid Eurocentrism. Indeed, this stance shouldn't be at all surprising since the main outlines of this very same understanding of social evolution as a process of moral-political learning that is characterized by greater degrees of reflexivity and decentration of worldviews go all the way back to Habermas's work on social evolution from the 1970s (Allen, 2016, p. 72f.).

In short, Allen's critique of Habermas has as its central point his theory of social evolution and his theory of modernity, that is, rationalization as a reconstruction of Western culture, while in the texts on cosmopolitanism Eurocentrism would manifest itself from the civilizing role attributed, according to her, to Western institutions and political culture.

2. Evaluating Allen's critique

The crux of Allen's critique appears to see rationality as the demands for a high degree of reflexivity and the demand for the presentation of reasons, arguments, and justification. According to her, such demands constitute a specific type of rationality that is tied to the West. This position becomes even clearer in her critique of Forst (whose theory is even closer to the Kantian tradition). In analyzing Forst's work and his reformulation of practical reason in terms of the right to justification, Allen asks:

> Can we be confident that "our" conceptions of practical reason are free of ideological distortions? Nor should this be seen as an empty worry, since there has been a great deal of criticism over the last thirty years or more, from feminist, queer, postcolonial, and critical race theorists, of the very conception of practical reason on which Forst's moral constructivism rests. Such critiques claim that the Kantian Enlightenment conception of practical reason explicitly or implicitly excludes, represses, or dominates all that is associated with the so-called *Other of reason*, whether that be understood in terms of madness, irrationality, the emotions, the affects, embodiment, or the imagination, all of which are symbolically associated with black, queer, female, colonized, and subaltern subjects. These symbolic associations serve both to rationalize and to justify existing relations of racial, heterosexist, and ethnic oppression and domination—by defining women, blacks, queers, colonized, and subaltern peoples as not rational and therefore as not fully human—at the same time that they reinforce certain stereotypical understandings of black, queer, feminine, and subaltern identity as closer to nature, more tied to the body, more emotional, more prone to madness, irrationality, and violence, and so on (Allen, 2016, p. 137f.).

In this passage the opposition rationality versus irrationality becomes clear. Reason, or rationality, invoked in both Habermas' and Forst's works as criteria for the acquisition of knowledge or for political decision-making processes, is argued by Allen to be a requirement that disempowers subaltern groups. Rationality is mobilized as a Western criterion of the disempowerment and oppression of groups who are described as unable to produce their knowledge or articulate their claims according to such criteria. Irrationality, on the other hand, is linked to emotions, madness, violence, and characteristics that modern societies attribute to stereotyped groups.

Allen's argument appears inclined to invalidate the distinction between rationality and irrationality, or even to deny the legitimacy of rationality as a criterion for evaluating knowledge or political discourses and actions on the grounds that such criteria are Western and neocolonial. Thus, reason (*voice*) is defined as the opposite of emotion (*noise*), and the former is invoked as a criterion for treating other forms of knowledge production or political manifestations pejoratively.

Allen's skepticism of any kind of practical reason is reaffirmed in her dialogue with Forst, when she maintains that "all our previous conceptions of practical reason have been exclusionary and dominating" and this "should undermine our confidence that our own conception of practical reason actually succeeds in transcending such entanglements" (Allen, 2019, p. 149).

One of the main points of divergence in this debate is the question of normativity. While Forst and Habermas continue to ground their theories in distinct versions of practical reason, Allen considers that this approach is foundationalist and imposes a legacy of Western modernity on other societies and cultures. She argues that the normativity of a theory that takes diversity seriously and avoids imposing itself in a colonialist manner should take the form of what she calls "metanormative contextualism", according to which:

> We could understand ourselves, at a first-order, substantive normative level, to be committed to the values of freedom, equality, and solidarity with the suffering of others, but understand these commitments, at the metanormative level, to be justified immanently and contextually, via an appeal to specific historical context rather than via an appeal to their putatively context-transcendent character. Such a metanormative contextualism offers a better way of instantiating the virtues of humility and modesty that are required for a genuine openness to otherness (Allen, 2016, p. 211).

Thus, she holds that it is possible for a non-relativist contextualism to coexist with a non-foundationalist universalism. Metanormative contextualism means, then, that the guiding principles are not grounded in a transcendental perspective (such as practical reason), but rather in a historical and contextualist one. In her view, contextualism is not the opposite of universalism since normative commitments considered universal in the scope of their application can be contextually grounded or justified. The opposite of contextualism is foundationalism (cf. Allen, 2019, p. 148f).

After presenting Allen's main arguments against practical reason (or communicative rationality in Habermas's case), two questions can be raised. First of all, to what extent does Allen's criticism actually affect the emancipatory potential of Habermas's theory, whether in domestic or international politics? Secondly, what are the consequences of adopting this criticism with regard to practical reason and its derivations (communicative rationality, public reason, or the right to justification)? Habermas's political theory offers some answers to Allen, which may reaffirm the importance of an ideal of practical reason, such as communicative rationality, for democratic and emancipatory practices.

Critiques of the theory of modernity: Allen posits that Habermas has adhered to the thesis of "multiple modernities", according to which

> the great world religions have had a great culture-forming power over the centuries, and they have not yet entirely lost this power. As in the West, these 'strong' traditions paved the way in East Asia, in the Middle East, and even in Africa or the development of cultural structures that confront each other today – for example, in the dispute over the right interpretation of human rights. Our Western self-understanding of modernity emerged from the confrontation with our own traditions. The same dialectic between tradition and modernity repeats itself today in other parts of the world. (...) The West is one participant among others, and all participants must be willing to be enlightened by others about their respective blind spots. (Habermas, 2010, p. 2)

This is where Habermas recognizes that modernity is defined by a process of confrontation between traditions within each society. It is a dialectic between tradition and reflections on tradition. He states, "[t]he West is one participant among others, and all participants must be willing to be enlightened by others about their respective blind spots" (Habermas, 2010, p. 2). The West's contribution to this process would be a kind of reflexivity that led to the secularization not of the private spheres of society, but of the public sphere. He is dealing here with changes such as the separation of law and morality or the methodology that gave rise to modern science, which he considers achievements that should not be lost (cf. Habermas, 2010, p. 3). Unlike Allen, I do not interpret these passages as constituting Habermas's reaffirmation of a new submission of other cultures to the Western one. Instead, I see it as a recognition of a specific contribution of a particular society to others. This thesis is restated in his 2019 book, *Auch eine Geschichte der Philosophie*, where he states that modernity should be understood as an arena in which different civilizations meet with more or fewer cultural specificities to give shape to their common infrastructure (cf. Habermas, 2019, p. 219).[4]

Critiques of cosmopolitanism: Allen argues that by considering the institutions of international law to be civilizatory, Habermas imposes Western institutions on other peoples and states considered uncivilized. In my understanding, however, Habermas's cosmopolitan project expresses the same openness to cultural plurality that was manifested in relation to modernity. What Habermas considers civilizing, since the end of the Second World War, has been the procedure of elaborating international law and building transnational and global institutions. It is by way of this procedure that involved parties can present their reasons and arguments, as well as their specificities, and perhaps avoid international relationships being determined merely on the basis of war and economic power. Two aspects of his analysis of international law support this interpretation.

He firstly points out that the United Nations itself has taken on an "inclusion orientation" by opening itself up to membership by any State willing to comply with the principles of the Universal Declaration of Human Rights rather than only to those States configured as liberal democracies. This openness to inclusion certainly generates contradictions, since some states formally adhere to the principles of the Universal Declaration of Human Rights, but do not follow them in practice. In the face of this, Habermas believes that tolerating contradictions between principles and practices can open a dialog between countries of the West and the East, with greater sensitivity to cultural, ethnic, and religious differences (cf. Habermas, 2006, p. 165f.).

4 In a recently published book entitled "A Critical Theory of Global Justice", Malte Frøslee Ibsen also considers that adherence to the paradigm of multiple modernities removes from Habermasian theory the label of culturalist and Eurocentrist that Allen seeks to attribute to it. He argues, "With this proposal, Habermas rejects the 'culturalist' image of world civilizations as incommensurable and mutually enclosed entities in favour of a more complex picture, where 'the great civilisations' are characterized by 'family resemblances', 'which imply similar cognitive developmental paths.' Accordingly, the view that contrasts 'European modernity' with temporally coexisting 'traditional' or 'premodern' forms of life in non-Western cultures, and which therefore necessarily regards the latter as developmentally inferior, is one that Allen imputes to Habermas, rather than a view that she identifies within his work. In other words, it is her assumption, rather than his" (Ibsen, 2023, p. 217). Another kind of critique of Allen's theory can be found in: Lopes, 2018.

Secondly, by pointing to the constitutionalization of international law as a rationalizing and civilizing process, Habermas emphasizes that the civilizing potential focuses on the procedures of international law-making rather than on substantive elements such as worldviews or ethical-cultural aspects. His critique of US actions after the 9/11 attack is a good example of the distinction between civilizing procedures and practices based on worldviews that, in turn, can be considered colonialist. He argued that the US government had violated the rules of international law and acted according to its own conceptions of right and wrong, overlapping the ongoing process of constitutionalization with a kind of moralization or ethnicization in the field of international relations. Thus, in the realm of international politics, a counter position emerges between law and morality — morality here understood as the ethical values proper to the American people (for example, the widespread notion of the struggle against evil in which the meaning of good and evil is attributed by the interested parties themselves). Thus, a people's own normative justifications replace procedures prescribed by international law. According to Habermas, "[e]ven an ultra-modern power like the US relapses into the pseudo-universalism of the ancient empires when it substitutes morality and ethics for positive law in issues of international justice." (cf. Habermas, 2006, p. 179-182).[5]

Critiques of communicative rationality in the theory of law and democracy: Finally, communicative rationality, as a type of practical reason, should not be considered an illegitimate or discriminatory criterion for evaluating social and political practices, as Allen argues. This criterion proves to be a powerful tool precisely of those groups that Allen calls "subaltern" (women and blacks, among others). An excellent example of this is Habermas's reflections on intolerance and discrimination. He argues that the issue of tolerance only arises after the prejudices that cause discrimination have been eliminated (such as sexism, racism, homophobia, xenophobia, etc.). It is important here to understand what characterizes some views as prejudice and discrimination. Habermas holds that it is the egalitarian and universalistic standards inherent in democratic citizenship that authorize considering some positions as discriminatory and prejudiced when they violate the equal rights of others and cannot be rationally justified. What is tolerable is that which is within reasonable disagreement; what is intolerable is that which discriminates and oppresses. According to Habermas, "the very same normative base of the constitution that justifies cultural rights and entitlements likewise limits a kind of aggressive self-assertion that leads to fragmenting the larger community" (Habermas, 2004, p. 18).

In his most recent work *on deliberative democracy and the new transformations in the public sphere* produced by digital media (Facebook, Twitter, Instagram, YouTube, etc.), Habermas is concerned about this lack of space for the presentation of arguments and reasons within the new media. These new environments can modify the public sphere to the point where it ceases to be "a sounding board for problems that must be processed by the political system" (Habermas, 1996, p. 359). He identifies one of the main transformations of the public sphere produced by the new media as being the difficulty in fulfilling the emancipatory promises of this type of communication. It has become more difficult to make digital platforms liable, even less so those users who express themselves in oppressive or discriminatory manners. This is due to the fact that these media are open to both progressive and reactionary agendas and Fake News.

3. Final Remarks

Within political philosophy and theories of democracy, reason and rationality have come to be challenged as elitist. This criticism is generally directed at theories of deliberative democracy that present very demanding normative criteria regarding argumentation and

5 A more detailed analysis of Habermas and international politics can be found in: Consani, 2020; Consani, 2021a.

justification. Moreover, the demands for the presentation of rational arguments in the public space are accused of preventing a vibrant democracy from flourishing by establishing a moral, and therefore pre-political, dividing line between *we* (*rational*) and *them* (irrational) and concealing the deep conflicts in today's societies (this is, for example, Mouffe's criticism of Rawls - cf. Mouffe, 2009).[6]

It is crucial to consider the consequences of this proposal. It is worth asking what might the potential implications be of eliminating the demands for rationality in the realms of science, morality, and politics. Recent events during the COVID-19 pandemic, such as the anti-vaccine movements and the defense of drugs without proven effectiveness (like hydroxychloroquine) remind us of the importance of not so quickly disregarding the requirements of modern science. In the political realm, on the other hand, the emotive and affective claims are not, by far, limited to subaltern groups. Oftentimes said claims have been raised by dominant groups, such as Trump or Bolsonaro supporters, who are not merely anti-science and anti-intellectual, but also racist and misogynistic. Those groups identified by Allen as subaltern (women, blacks, queers, indigenous people, etc.) are often those who have presented their claims with great propriety, making use of consistent argumentation and justification that is completely compatible with the requirements of practical reason.

Forst has correctly pinpointed a grave error underlying Allen's arguments. He argues that it is in fact the decolonial critics who disrespect and infantilize subaltern groups by considering them discursively incompetent or incapable of demanding or offering justifications. He further identifies a serious methodological error in this critique, which is the confusion between the descriptive and normative spheres (genealogy and validity). He states that the demanding structure of justification is part of the grammar of social conflict that belongs to emancipatory struggles and is not tied to a specific way of life (European or Western). In his words: "The language of emancipation and of no longer wanting to be denied one's right to be a participatory equal is a universal language spoken in many tongues" (Forst, 2014, p. 184).

Forst does not agree with Allen's characterization that "all our previous conceptions of practical reason have been exclusionary and dominating" (Allen, 2019, p. 149). This critique, he says, is "too general and sweeping", since it minimizes the effective performance of rational critique in challenging unreasonable forms of domination (religious, racial, gendered, or nationalistic). Thus, he states that concepts of rationality tend to generate domination when they affirm the context, rather than when they question it. He also warns that the proper critique of universalism, namely that which rejects false universals[7], cannot be compared to support for parochialism (cf. Forst, 2019, p. 182).

Allen contradicts herself, says Forst, when she argues for the existence of a universalist first normative order and a contextualist second one. This would require that the validity of universalist claims depended upon contextualist interpretations, that is, on social and historical standpoints. This would undermine the critical aspect of the theory in relation to the dominations of its time and context. For Forst, a universalist contextualism is nothing but a false universalism, i.e., a relativist proposition (cf. Forst, 2019, p. 183).[8]

6 On this topic see Consani, 2021b.

7 As highlighted above, an example of the unveiling of false universalism is found in Habermas's work on cosmopolitanism when he criticizes the role played by the USA on the international scene after the 9/11 attack.

8 In the follow-up to this debate with Allen, Forst defends his theory and makes the case for Allen's proposed methodology as follows: "(...) I regard humans as contextual beings, but I believe that their capacity to reason critically enables them to transcend their horizons and to achieve what Gadamer called a 'fusion of horizons'. Morover, I share the Kantian view that reconstructing the rational criteria for justifying validity claims is the work of *finite* beings who engage in an exercise of reconstruction that they perform from within their local and finite perspectives, for want of any other perspective; they think that they have a faculty of communication and reasoning that is governed by principles that they can use for the purposes of knowledge and critique. To regard the reconstruction of these principles as a tran-

Forst's critique of Allen highlights the risks of losing rationality as a criterion for evaluating social action (moral or political). In the absence of criteria that go beyond context and scrutinize emotion, it seems difficult to distinguish legitimate actions from those that seek to perpetuate relations of domination. In recognizing the relevant role played by conceptions of practical reason as criteria for the legitimation of social and political actions, I do not intend in any way to disregard the importance of emotion and affect in the social or political field. It is essential, however, to distinguish between *comprehension* and *justification*.[9] Emotions (affections, identifications, resentments, hatred, etc.) are a social and political force and are often the first form of expression or manifestation of subjective perspectives regarding relations of domination. Emotions certainly "sound the alarm" regarding social injustice, and failing to understand their role would be to neglect a crucial aspect of society and politics. However, without a rational, objective, and non-contextualist criterion that can support the justification of standards of legitimacy, it may be difficult to differentiate social claims that seek to break with relations of domination (Black Lives Matter, for example) from those that seek to maintain them (US Capitol attack or a mob of Bolsonaro's supporters who attacked Brazil's federal government buildings, for instance).

Bibliographic References

ALLEN, A. *The end of progress: decolonizing the normative foundations of critical theory.* New York: Columbia University Press, 2016.

ALLEN, A. Progress, Normativity, and Universality: Reply to Forst. In: Amy Allen and Eduardo Mendieta (ed.). *Justification and Emancipation: The Critical Theory of Rainer Forst.* Pennsylvania: The Pennsylvania University Press, 2019, p. 145-155.

CONSANI, C. F. É possível uma Constituição sem Estado? Uma análise da proposta habermasiana de política mundial sem governo mundial. *ETHIC@* (UFSC), 19, p. 318-337, 2020.

CONSANI, C. F. A constitution without a state? An analysis of the Habermasian proposal for global politics without a world government. In: Cristina Consani, Joel T. Klein and Soraya Sckell (Org.). *Cosmopolitanism: From the Kantian Legacy to Contemporary Approaches*. Berlin: Duncker & Humblot, p. 189-202, 2021a.

CONSANI, C. F. Consenso, conflito e o conceito do político: uma avaliação das críticas de Mouffe a Rawls. In: Cristina Foroni Consani. Julia Sichieri Moura and Nythamar de Oliveira (Org.). *Justiça e Libertação: A Tribute to John Rawls*. Porto Alegre: Fundação Fênix, p. 137-161, 2021.

EKLUNDH, E. *Emotions, Protest, Democracy. Collective Identities in Contemporary Spain.* London/

scendental enterprise in no way turns my approach into a form of foundationalism, since (in contrast to metanormative contextualism) it doesn't claim to possess a God's-eye view but only to perform a finite rational reconstruction. Thus, the reconstruction can be fallible, but we can recognize this fallibility only if we believe that the principles of reason are reliable and binding for us, especially when overcoming our failures – to believe otherwise would render our fallibility unintelligible. Kantians are well able to distinguish between a *transcendental* argument and an absolute *transcendent* standpoint detached from finitude and context. Unlike metanormative contextualism, I have no access to a 'genuinely context-transcendent point of view' (Forst, 2019, p. 184).

9 This distinction and its impacts on the relationship between reason and emotion in theories of democracy will be explored in a further article.

New York: Routledge, 2019.

FORST, R. Justifying Justification: Reply to My Critics. In: Davod Owen (ed.) *Justice, Democracy and the Right to Justification: Rainer Forst in Dialogue*. London: Bloomsbury, 2014.

FORST, R. The Justification of Progress and the Progress of Justification. In: Amy Allen, Eduardo Mendieta (eds). *Justification and Emancipation: The Critical Theory of Rainer Forst*. Pennsylvania: The Pennsylvania University Press, p. 17-37, 2019.

FORST, R. Navigation a World of Conflict and Power: Reply to Critics. In: Amy Allen, Eduardo Mendieta (eds). *Justification and Emancipation: The Critical Theory of Rainer Forst*. Pennsylvania: The Pennsylvania University Press, p. 157-187, 2019.

HABERMAS, J. *Communication and the Evolution of Society*. Translated by Thomas McCarthy. Boston: Beacon, 1976.

HABERMAS, J. *The Theory of Communicative Action*. Vol. 1, *Reason and the Rationalization of Society*. Translated by Thomas McCarthy. Boston: Beacon, 1984.

HABERMAS, J. *The Theory of Communicative Action*. Vol. 2, *Lifeworld and System: A Critique of Functionalist Reason*. Translated by Thomas McCarthy. Boston: Beacon, 1987.

HABERMAS, J. Discourse Ethics: Notes on a Program of Philosophical Justification. In *Moral Consciousness and Communicative Action*, translated by Christian Lenhardt and Shierry Weber Nicholsen. Cambridge, Mass.: MIT Press, 1990.

HABERMAS, J. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Translated by William Rehg. Cambridge, Mass.: MIT Press, 1996.

HABERMAS, J. Intolerance and discrimination. *Oxford University Press and New York University School of Law*, Oxford; Nova York, 2003, 1 (1), p. 2-11.

HABERMAS, J. Religious tolerance: the pacemaker for cultural rights. *Philosophy*, Cambridge, 2004, 1, p. 5-18.

HABERMAS, J. *The Divided West*. Edited and translated by Ciaran Cronin. Cambridge: Polity, 2006.

HABERMAS, J. *A Postsecular World Society? On the Philosophical Significance of Postsecular Consciousness and the Multicultural World Society*. Interview with Eduardo Mendieta. http://blogs.ssrc.org/tif/wp-content /uploads/2010/02/A-Postsecular-World-Society-TIF.pdf.

HABERMAS, J. Interview with Jürgen Habermas. In: Bächtiger, Andre et al. (Ed.). *The Oxford Handbook of Deliberative Democracy*. Oxford: Oxford University Press, 2018, p. 872-881.

HABERMAS, J. *Auch eine Geschichte der Philosophie*. Berlin: Suhrkamp, 2019.

HABERMAS, J. Reflections and Hypotheses on a Further Structural Transformation of the Political Public Sphere. *Theory, Culture & Society*, 2022, 39(4), p. 145–171.

IBSEN, M. F.. *A Critical Theory of Global Justice*: The Frankfurt School and World Society. New York: Oxford University Press, 2023.

LOPES, A. C. O Que é Crítico na Descolonização da Teoria Crítica? Amy Allen e O Fim do Progresso. *Dissonância*: Revista de Teoria Crítica, 2, 2018, p. 66-89.

MENDIETA, E. From imperial to dialogical cosmopolitanism. *Ethics & Global Politics*, 2(3), 2009, p. 241-258.

MIGNOLO, W. D. The many faces of cosmo-polis: Border thinking and critical cosmopolitanism. *Public Culture*, 12 (3), 2000, p. 721-748.

MIGNOLO, W. D. Cosmopolitanism and the De-colonial Option. *Stud Philos Educ*, 2010, 29, p. 111-127.

MIGNOLO, W. D. *The politics of decolonial investigations*. Durham and London: Duke University Press, 2021.

MIGNOLO, W. D; WALSH, C. E. *On Decoloniality: Concepts, Analytics, Praxis*. Durham and London: Duke University Press, 2018.

MOUFFE, C. The limits of John Rawls Pluralism. In: *Theoria: A Journal of Social and Political Theory*. 56 (118), 2009, pp. 1-14.

SANTOS, L. R. dos. *Kant Pensador da Terra*. Lisboa: Edições Colibri, 2022.

Recebido 12/2022 | Aprovado 03/2023

Resenha: MAREY, Macarena. Voluntad omnilateral y finitud de la Tierra. Una lectura de la filosofía política de Kant

[Argentina, Adrogué: La Cebra, 2021, 336 pp.]

Luciana Soria Rico

Universidad de la República de Uruguay (Montevideo, Uruguai)

DOI: 10.5380/sk.v21i1.91544

O livro de Macarena Marey (Universidad de Buenos Aires) é fruto de uma longa trajetória de investigação na qual a filósofa argentina desenvolve uma exegese sistemática da filosofia política de Immanuel Kant e mostra suas ressonâncias na "atualidade filosófica" (Marey, 2021, p. 26).

A obra de Marey percorre uma perspectiva histórica dupla na qual, por um lado, põe ao autor em diálogo com outros referentes da filosofia política moderna tais como Jean Jacques Rousseau e Thomas Hobbes, expondo influências e discrepâncias teóricas que influenciaram ao filósofo alemão a recolher preocupações anteriores e reformular e sistematizar conceitos. Por outro lado, o livro revela como algumas interpretações kantianas contemporâneas orientadas por John Rawls e Jürgen Habermas, amplamente difundidas e aceitas na esfera intelectual e acadêmica, deturparam ou desvirtuaram teses centrais de sua obra político-filosófica. Parte dessas fragilidades se deve ao fato de que ou eles ignoraram os escritos políticos relevantes de Kant, ou suas interpretações se basearam em uma análise reduzida de pequenos escritos da década de 1780, mal conhecendo os da década de 1790, e por isso reputaram suas concepções políticas à ética. Na década após a Revolução Francesa, Marey compreende que foram publicadas obras-chave – como *Metafísica dos Costumes*, *Religião nos Limites da Razão* e *Concurso das Faculdades* – nas quais o filósofo de Königsberg não só deu uma guinada conceitual, mas amadureceu muito suas ideias oriundas da década anterior. Nesse sentido, a autora dedica o livro à reconstrução de ideias centrais do *corpus* político kantiano ao apresentar suas próprias teses sobre a compreensão da normatividade subjacente.

Em conciliação com o exposto, o livro aborda a relação entre ética e direito, primeiro, apresentando uma série de razões para rejeitar a forma onilateral em que costuma ser apresentada a relação entre os dois e, segundo, mostrando diretamente a impossibilidade da derivação do direito a partir da ética. Defende-se que a ética e o direito respondem a diferentes necessidades normativas e que, portanto, nenhum pode cumprir a função do outro ou substituí-lo. Para apoiar isso, Marey oferece certas razões pelas quais Kant pensa que o direito deve ser moral e "uma razão filosófico-política pela qual o direito não pode ser ético" (Marey, 2021, p. 41). Isso estará relacionado à ideia de uma comunidade ética kantiana que só pode ser realizada dentro de comunidades políticas onde os sujeitos adquirem direitos e obrigações correlatas, não fora delas. Isso leva Marey a afirmar que o critério normativo kantiano prioriza o político como realização plena do ético. Tal prioridade não exclui que a comunidade ética possa então atuar na esfera política já formada – onde o jurídico e o legal não bastam – e como a ética da virtude tem o poder de mobilizar a ética comunitária contra as injustiças estruturais.

Nessa linha, argumenta-se que uma das tarefas que o direito ocupa em Kant é explicar como as

obrigações jurídicas, que por natureza não existem, surgem e se sustentam e "cuja geração não implica uma imposição onilateral" (Marey, 2021, p. 121). Marey desenvolve sua própria tese interpretativa na qual fundamenta que a ideia de igualdade inata é a fonte da normatividade jurídica kantiana que finalmente confere o caráter moral ao direito. Nesse ponto, ela argumenta que Kant conecta obrigações jurídicas externas – baseadas na posse do que é *meu propriamente* – com a necessidade moral do Estado por meio de uma característica intrínseca da própria obrigação: a reciprocidade universal. Qualquer obrigação legal que aspire à legitimidade normativa tem que satisfazer essas condições de reciprocidade universal, só que há coação legítima, caso contrário haveria mera imposição onilateral sem força. Com isso, percebe que a única fonte de onde emana a obrigação kantiana é a reciprocidade universal e só é realizável em um estado jurídico. E isso leva a perceber o fundamento moral das vontades investidas de autoridade política no *exeundum* particular kantiano.[1]

De acordo com o exposto anteriormente, o livro trata da resolução kantiana sobre a conformação da unidade do corpo político. Ao contrário de outros contratualistas, Kant não pretende fundar o Estado sobre uma razão instrumental-prudencial, mas moral-prática, e para isso fará um novo uso da categoria de soberania popular, buscando evitar cair em visões agregativas ou multiplicidade irredutível das vontades.

Marey argumenta que a criatividade kantiana para resolver isso está em conceber a soberania popular como uma vontade onilateral e unificada a priori na qual o povo é sustentado pela ideia de um direito válido e legítimo. A vontade entendida onilateralmente refere-se à reciprocidade de obrigações e direitos-deveres de correspondência, é a fonte do direito validamente vinculante e, portanto, é a condição jurídica que possibilita a aquisição de direitos coercitivos e recíprocos e obrigações correlacionadas. Nesse sentido, Marey sustenta que a vontade geral do povo não é pura idealidade, mas um critério normativo para avaliar instituições ou poderes de fato instituídos local ou transaccionalmente. Argumenta-se que a soberania popular kantiana é transversal às três esferas do direito público (Constitucional, Do povo e Cosmopolita) que compartilham sua justificativa normativa de equidade e justiça. Essas três esferas do direito repousam na ideia de posse comum da terra e na tese correlativa de que somente as vontades onilaterais podem criar leis legítimas.

Isso nos leva a outra das questões relevantes que ocupam o livro, esta é: a discussão sobre a natureza provisória dos direitos de propriedade e sua relação com a posse comum da terra. Embora esta já fosse uma premissa teológica da filosofia medieval, Marey aponta que em Kant ela possui pelo menos duas características que a tornam original. Uma delas é que a posse comum da terra é uma ideia metafísica e racional que contém a priori a possibilidade de posse, mas apenas na medida em que está ligada à ideia de uma vontade onilateral.

E segundo, que diferentemente da concepção teológica de Kant, ela não sustenta que Deus nos deu a terra em comum como uma herança compartilhada para preservação, mas sim que, porque a terra é finita, os homens são obrigados a coexistir como vizinhos de uma comunidade cosmopolita que nos leva a afetar uns aos outros em nossas liberdades externas, e à necessidade de regular os direitos de posse. Nessa linha, para Kant não haveria direitos de propriedade pré-políticos, mas sim estabelecidos de forma recíproca e iníquo como tal em uma comunidade política. O *seu* e o *meu* adquirem status legal somente quando os direitos e deveres de propriedade são regulados reciprocamente sob a consideração da finitude da terra. Os direitos de propriedade adquiridos são então legítimos quando passam por esse processo, razão pela qual Marey argumenta que Kant fornece um critério crítico com objetividade prática para avaliar a legitimidade das relações de propriedade existentes, que são sempre provisórias, já que nunca são direitos de propriedade no sentido exclusivo e absoluto. A autora destaca a conexão analítica entre a vontade onilateral e a finitude da terra, dando o título ao livro.

1 Esse princípio da filosofia prática kantiana funciona como um imperativo categórico que afirma o dever jurídico de sair do estado de natureza. A partir dela, o Estado se funda na ideia (Marey, 2021, p. 93).

Finalmente, gostaria de mencionar dois aspectos do livro que são independentes de suas contribuições conceituais substantivas. Em primeiro lugar, sua visão crítica das leituras e interpretações filosóficas que poderiam se estender para além do autor analisado, neste ponto, destacaria a rejeição da pretensão de desvendar um filósofo-em-si (no caso "um Kant-em-si"), ou seja, a busca de uma interpretação final e autêntica das teses fundamentais (no caso da filosofia prática kantiana), por uma vez encontrada limitada à uma exposição de museu. Marey entende que a interpretação filosófica é uma tarefa em aberto, e por isso a revisão do arquivo filosófico kantiano não deve ser esgotada. Por outro lado, independentemente das contribuições específicas que este trabalho deliberadamente pretende, também há lições e novos significados para a filosofia política para além de Kant, entre eles destacaria o fundamento sobre por que os critérios normativos não são necessariamente ideais e se pretendem ser críticos, eles não podem ser; e a crítica dos direitos entendidos onilateralmente pelo liberalismo. Essas e outras contribuições do livro o tornam interessante como um texto de filosofia política em si.

Em suma, o livro de Marey oferece leituras interpretativas alternativas em relação às amplamente estendidas, sem sustentar a pretensão de um "acesso privilegiado às verdades exegéticas" (Marey, 2021, p. 36), mas sem abandonar a pretensão da rigorosidade conceitual para implantar uma leitura própria do sustento fundado. Por fim, destaco do livro não apenas suas interpretações inéditas para especialistas, mas também sua grande clareza expositiva, apesar de não ser um livro de divulgação pretendida, em minha opinião, sua leitura é acessível a não especialistas em estudos kantianos. Em síntese, a publicação de Marey em espanhol pela editora argentina La Cebra enriquece os estudos kantianos, razão pela qual é uma leitura recomendada para todos os interessados em conhecer ou aprofundar a filosofia política kantiana.

Referências Bibliográficas

MAREY, M. *Voluntad omnilateral y finitud de la Tierra. Una lectura de la filosofía política de Kant*. Buenos Aires, La Cebra, 2021.

Recebido 11/2022 | Aprovado 02/2023

Reseña: Luciana Martínez, Esteban Ponce (Eds.), El genio en el siglo XVIII

[Barcelona: Herder, 2022]

Martín Iván Druvetta[1]

Universidad Nacional de Rosario (Rosario, Argentina)

DOI: 10.5380/sk.v21i1.91986

El genio en el siglo XVIII es un libro reciente, del 2022, publicado en lengua española por la editorial Herder, en el que participa un grupo de investigadores especialistas en el área de filosofía moderna. El libro, editado por Luciana Martínez y Esteban Ponce, aborda la figura del genio desde distintos ángulos. Para ello, propone un análisis histórico y conceptual desde los inicios de la Francia pre-ilustrada hasta el auge del Romanticismo alemán (cap. 1-12), una revisión del término atendiendo a su significado en la lengua española, poniendo el foco en el siglo XVIII hispanoamericano (cap. 13), y un estudio desde la óptica de dos mujeres, lo cual habilita acertadamente una lectura no androcéntrica del término (cap. 14-15).

El libro ofrece diferentes acercamientos acerca de la figura del genio, y si bien proporciona concepciones un tanto diversas, logra exhibir su carácter complementario. Sus dieciocho capítulos (cada uno escrito por un autor distinto) se encuentran muy bien articulados, de modo que, atendiendo a los distintos elementos condicionantes de cada contexto histórico, el libro consuma una idea muy completa del *genius*.

En el primer capítulo, Nicolás Olszevicki proporciona una reconstrucción general e histórica del genio, reparando en distintos elementos históricos de la Francia pre-ilustrada. Se trata de un apartado que repara principalmente en la génesis del concepto. En el segundo capítulo, Kamila Babiuki, centrándose en J.-B. Dubos (1670-1742), principalmente en su obra *Reflexión crítica sobre la poesía y sobre la pintura* (1717), vislumbra ciertos rasgos innatos en la propia naturaleza del genio, lo cual supone la idea de una fisiología única que lo predispone a ser como tal. En el tercer capítulo, Esteban Ponce marca que, según D. Diderot (1713-1784), existe algo que no puede ser comunicado ni enseñado en la propia genialidad, pero esto no quita que sea posible analizarlo. La naturaleza del genio pertenece a ese mundo materialista y no mecanicista diderotiano, el cual, si bien no es posible aprehenderlo, sí es posible establecer cuáles son las condiciones que lo determinan.

La segunda parte del libro se ocupa del pensamiento ilustrado inglés. En ella, en el cuarto capítulo, Luis F.S. Nascimiento relaciona la "genialidad" y la "filosofía" en la Inglaterra del siglo XVIII. Para ello, primero, parte de la definición que J. F. de Saint-Lambert (1716-1803) toma de la *Encyclopédie* (1751) de Diderot y D'Alembert (1717-1783); luego, mediante la oposición de varios conceptos, como "belleza" y "verdad", "razón" y "sensibilidad" e "ingenio" y "juicio", contrasta la concepción de genio desde la óptica de varios autores, como J. Locke (1632-1704), A. A. C. Shaftesbury (1671-1713), J. Addison (1672-1719) y E. Young (1683-1765).

1 martindruvetta@gmail.com

En el quinto capítulo, Valeria Schuster se apoya en D. Hume (1711-1776) para delinear el concepto de genio. Si bien la imaginación, como aquella que determina el contenido de las representaciones, ha sido ampliamente trabajada en la literatura humeana, poco se ha dicho sobre su uso creativo. En este punto se focaliza el análisis propuesto por Schuster, quien identifica que hacia finales del siglo XVII, la palabra *genius* tenía al menos seis significados, una pluralidad semántica que incluso fue aumentando avanzado el siglo XVIII. Finalmente, Schuster marca la relación que Hume establece entre el artista y el crítico, como así también el vínculo entre la genialidad y el placer, entendida la genialidad como una práctica placentera.

En el sexto capítulo, Alexandre Amaral Rodrigues, centrándose en Alexander Gerard (1728-1795), principalmente en su obra *An Essay on Taste* (1759), subraya la invención en el genio y su lazo con el arte bello, marcando la reciprocidad existente entre la imaginación y la razón. El autor analiza la forma en que la razón asiste a la imaginación.

Luego, el libro se ocupa de la filosofía ilustrada alemana. En el séptimo capítulo, Julio del Valle aborda la noción de genio a partir del autor A. G. Baumgarten (1714-1762), marcando la importancia que tuvo este concepto en el proceso secular llevado a cabo por la Ilustración. Para ello, distingue entre *genius* e *ingenium*, y sostiene que en tal distinción se haya la génesis del genio como creador.

En el octavo capítulo, Luciana Martínez se focaliza en I. Kant (1724-1804), puntualizando principalmente en *Kritik der Urteilskraft* (1790), aunque repara también en el llamado *corpus* antropológico kantiano. Martínez se propone delinear el concepto de genio propuesto por Kant, el cual puede definirse a la manera de un don provisto por la naturaleza (único e incapaz de adquirirse), necesario para el arte e incapaz de manifestarse en cuanto tal en ninguna otra actividad o práctica del ser humano que no sea la artística. Este recorrido le permite a Martínez marcar la particularidad anímica del genio, como así también (a partir de la distinción entre "espíritu" e "idea estática") lo peculiar y distintivo que es el arte respecto del resto de las producciones humanas.

En el noveno y en el décimo capítulo se analiza la figura del genio en el marco del romanticismo alemán. Por un lado, Virginia López Domínguez se centra en J. G. Herder (1744-1803), para establecer un vínculo entre la genialidad y la libertad, a la vez que marca cómo el culto al genio guarda un carácter político, reivindicando ciertas prácticas de la burguesía por encima de ciertas prácticas de la nobleza. Por el otro lado, María Verónica Galfione se centra en F. Schlegel (1772-1829), y sostiene que si bien se suele asociar el Romanticismo a una poética del genio, Schlegel, para quien dicho concepto no ocupa un rol central en su filosofía, parece quedarse fuera de la norma. Esto hace suponer a Gelfione que, en términos schlegeliano, el *genius* es sinónimo de talento, por lo que la capacidad reflexiva se halla por encima de la propia genialidad.

En el décimo primer capítulo, Miguel Alberti sostiene que para G. P. F. von Hardenberg (1772-1801), más conocido por su pseudónimo Novalis, el concepto de genio adopta una significación mucho mayor; el término, precisamente por su alcance semántico, se aleja de los usos corrientes de la época. De esta manera, Novalis visualiza que el terreno afín al genio no es únicamente el estético, sino también todo aquello que implique la creación del mundo que lo rodea. Esto hace que la genialidad no se limite a una actividad de penetración artística o invención del pensamiento, sino que se proyecte a toda la vida en general.

Por último, el libro contiene cuatro capítulos centrados en tradiciones o perspectivas filosóficas menos usualmente reconocidas. En el décimo segundo capítulo, Manuela Sanna analiza, desde una perspectiva filológica, el concepto de genio en la obra de G. Vico (1668-1744), y para ello propone un estudio minucioso del término *ingenium*, describiendo la proliferación lingüística que tuvo a lo largo de los años, como así también el vínculo filológico con la palabra *inventio*.

En el décimo tercer capítulo, Raúl Trejo Villalobos propone el examen de algunos significados de las palabras "genio" e "ingenio" en la lengua española. Para ello, divide su investigación en tres partes. Una primera que aborda la terminología de diccionarios del s. XX hasta llegar al *Diccionario de sinónimos castellano* de J. Justo de Gómez de Cortina (1799-1860) y a *Arca de letras y teatro universal* de J. A. de Navarrete (1749-1814). Una segunda en la cual examina la concepción kantiana del genio desarrollada en la *Kritik der Urteilskraft*. Una tercera en la que se focaliza en cuatro escritos: *Oración vindicativa del honor de las letras, y los literatos* (1763), de M. C. Coriche, *Sobre lo bello en general* (1801), de P. J. Márquez (1741-1820), *El lazarillo de ciegos caminantes* (1771-1773), de Calixto Bustamante Carlos Inca (1706-1783) y *El nuevo Luciano de Quito* (1779), de E. de Santa Cruz y Espejo (1747-1795). De esta manera, Villalobos afirma la existencia de dos aspectos posibles del genio en Hispanoamérica, los cuales responden a significados distintos, uno mitológico (siglo XVIII) y el otro antropológico (siglo XIX), culminando con cierta insinuación hacia las ideas del genio de Kant y Márquez.

En el décimo cuarto y en el décimo quinto capítulo, encontramos traducciones de textos escritos por mujeres: Mary Wollstonecraft y J. M. Le Prince de Beaumont (1711-1780). Estas traducciones se encuentran prologadas y anotadas por Kamila Babiuki y Mariela Paolucci y Natalia Zorrilla. En el texto de Beaumont se hace visible qué tan determinante pueden ser las condiciones materiales no sólo en la constitución del sujeto, sino en la propia genialidad. Todo el capítulo puede leerse como una crítica a la época por el rol que la misma sociedad le atribuía a la mujer. Wollstonecraft, por su parte, propone una idea del genio a partir de un análisis crítico de la imitación, comparando la belleza natural con la belleza artística.

El libro *El genio en el siglo XVIII* contribuye un gran aporte no sólo para el estudio de la estética sino para otras áreas de la filosofía. Precisamente, el estudio meticuloso de la investigación que lleva a cabo visualiza una noción del genio de la modernidad inmensamente compleja, a la vez que fundamental para la comprensión de todo el contexto filosófico de la época. Así también, hay que remarcar que los artículos brindan una narrativa ordenada y con gran claridad expositiva, lo cual permite hacerse de la idea que cada autor propone describir.

Referências Bibliográficas

MARTÍNEZ, L. & PONCE, E. (Eds.). *El genio en el siglo XVIII*. Barcelona: Herder, 2022.

Recebido 03/2023 | Aprovado 03/2023