UDK: 172.4 Kant I.
https://doi.org/10.2298/FID1804598C
Original Scientific Article
Received: 21.07.2018. Accepted: 12.08.2018.

PHILOSOPHY AND SOCIETY
VOL. 29, NO. 4, 471–646

Luigi Caranti

REPLIES: AUTONOMY AND HUMAN DIGNITY. A REASSESSMENT OF *KANT'S POLITICAL LEGACY. HUMAN RIGHTS, PEACE, PROGRESS*

ABSTRACT

The paper centers on some problematic theses of my book *Kant's Political Legacy. Human Rights, Peace, Progress* (UWP 2017). This reconsideration is occasioned partly by comments I received and partly by my own process of self-criticism. I focus on the point that commentators have mainly criticized, that is, the link I suggest between human dignity and our capacity for moral behavior, or autonomy. The first part recalls the basic features of my Kant-inspired and yet in many regards anti-Kantian account of the relation between dignity and autonomy and replies to some criticisms received from orthodox Kantians. The second part is strictly connected to the first because it deals with the reasons we have to believe that we are autonomous. While in the book I sketched Kant's own reasons for the 'reality of freedom,' as he puts it, I focus now on Bojan Kovačević's suggestion to look at characters in novels written by artistic geniuses (in particular Leo Tolstoy) to find indirect evidence in favor of autonomy. This allows me to reflect on the kind of evidence one can legitimately expect in the proof at issue. Thirdly, I reply to a classical objection, ignored in the book, that impacts with equal force Kant's ethics and my own position. The problem concerns people with temporary or permanent impairment of rational capacities. If I let human dignity depend on our capacity for autonomous behavior, am I committed to the counterintuitive (and rather devastating) conclusion that children or people suffering from momentary or irreversible loss of rational capacity, and *a fortiori* of autonomy, do not have dignity and therefore do not deserve to be protected by human rights?

KEYWORDS

morality, autonomy, dignity, human rights, Kant, mental impairment, speciesism

In my book *Kant's Political Legacy. Human Rights, Peace, Progress*, as the title indicates, I attempt to show how Kant's ideas can be illuminating for three themes crucial both for contemporary politics and for political theory. In particular, I sketch a foundation of human rights considerably different from any foundational argument available today in the literature. Further, I provide an interpretation of Kant's model for peace, that I show to be profoundly different from the one assumed by the majority of scholars and by the research program – the democratic (or liberal) peace – that made Kant's pacifism known and studied well beyond the circle of interpreters. Finally, in the most speculative part of the book I offer a defense of Kant's teleology, which I take quite seriously as meant to provide objective

Luigi Caranti: Professor of Political Philosophy, DSPS, Università di Catania; lcaranti@unict.it.

reasons to believe that progress towards a future of peace is more likely than any other outcome. Almost two years have passed since the book was published and I had the opportunity to rethink some of my theses, either through a process of endogenous reconsideration, or stimulated by the criticisms attracted.

Since commentators have mainly focused on the thesis, central for my foundation of human rights, that humans have dignity because they possess autonomy, I devote the paper to a reconsideration of this tenet. The first part deals with the legitimacy of my Kant-inspired and yet in many regards anti-Kantian reading of the relation between dignity and autonomy. I was aware that my idea that agency not inspired by the Categorical Imperative could count as authentically autonomous would attract criticisms by orthodox Kantians. I take here the opportunity to defend myself to the extent that this is possible. The second part is related to the first in that it deals with the reasons we have to believe that we are autonomous. While in the book I sketched Kant's own reasons for the 'reality of freedom,' as he puts it, especially in the second critique, I start now from Bojan Kovačević's suggestion to look at characters in novels written by artistic geniuses (in particular Leo Tolstoy) to find indirect evidence in favor of autonomy. This allows me to reflect on the kind of evidence one can legitimately expect in the proof at issue. Thirdly, I reply to a classical objection to Kant's ethics, that impacts just as deeply on my own reading, in a way that I had not done in the book. The problem concerns people with what I call temporarily or permanently impaired autonomy. If, with Kant, we let human dignity depend on our capacity for autonomous behavior, it seems that we are forced to the counterintuitive (and rather devastating) conclusion that children or people suffering from momentary or irreversible loss of rational capacity (and *a fortiori* of autonomy) have neither dignity nor human rights.

1. Autonomy and Human Dignity

The working hypothesis of *Kant's Political Legacy*, part I, is that all current foundational accounts of human rights lack a crucial ingredient. This is the simple intuition, latent in all major documents of human rights, that human beings are *worthy* creatures, despite the atrocities for which they have been responsible. Within 'humanity', by virtue of which, the documents say, we have human rights, there is supposed to be something extraordinarily valuable and awe-inspiring that makes it obligatory never to treat humans below a certain threshold of respect. This kernel of value, which a foundation should spell out, serves as protection not only against the violations of our dignity that others cause to us, but also against the degradation we can bring upon ourselves.

In starting from this working hypothesis, my approach goes in a direction opposed to the practical compromise dear to founding fathers of the culture of human rights. Jacques Maritain and Eleanor Roosevelt famously stated that we all agree on HR and human dignity "on the condition that no one asks us why." If a foundation is supposed to convince a skeptic (for what other reasons would you want a foundation?), I think we must say something as to why human beings have dignity, why they are to be esteemed independently of their religion, citizenship, gender, race. Only if we do so will we be in the position to counter the arguments

of those who believe that some, but not exactly all persons deserve the equal respect promised by HR. In choosing this path, I knew that my position would appear as a form of ultra-orthodoxy, or perhaps of 'foundationalism', to use Tasioulas' derogatory label (Tasioulas 2015: 46–47). To make things worse, in construing the required account of human dignity I took my lead from Immanuel Kant and even if I quickly depart from him on a number of crucial points, this choice is bound to be perceived as somewhat bizarre. Among other things, Kant's account of human dignity seems to rest on his controversial doctrine of our belonging not only to the phenomenal, but also to the noumenal world, a very metaphysical tenet indeed.

Still, a Kantian foundation should not be discarded before one critically explores its potential. Kant offers one of the most compelling accounts of our dignity in the history of thought, and this cannot be fully irrelevant for HR. Moreover, silence on the reasons we have to believe that human beings have dignity will not convince those who think that it is permissible to treat some of us below certain standards. If we do not know *why* we have dignity, it is very likely that we will not know what this entails, and we will have a difficult time clarifying why and when human rights trump normative considerations in the form of maximization of general utility, the furthering of certain ideologies, or the protection of traditional practices. In addition, failing to unpack the notion of human dignity gives no direction regarding related questions crucial for the theory of human rights. It is unlikely that we will know what rights should count as *human* rights unless we know why human rights should exist in the first place. It may not be impossible to come up with an answer to this and other similar questions before we clarify what dignity is and entails. But few would deny that some clarity on human dignity will help greatly in dealing with those questions.

If a Kant-inspired foundation of human rights is worth exploring, the following might be a reasonable way to proceed. The central thesis, in itself far from original, is that humans have dignity because they are capable of a unique form of freedom, namely autonomy, which allows them to perform actions motivated by what they take as morally obligatory. Autonomy is not to be understood merely as the ability to choose one's path in life, or as the ability to be rational in the sense of purposive agents. With Kant, we refer to a capacity distinct from and 'higher' than practical freedom. We have in mind the ability to act under self-imposed moral constraints.

The distinction between the common-sense notion of autonomy (self-determination) and the one at work here (capacity for moral agency) is important to understanding the way in which our approach links the possession of a faculty to the intrinsic worth that entitles humans to the protections of human rights. In fact, why should the sheer possession of a capacity ground any worth? The fact that humans have a peculiar capacity hardly grounds any special right or entitlement. Ours is probably the only species that kills for sheer amusement (cats may be another example) or for cold, long-term calculation of interest. It may also be the only one capable of lying with a high level of sophistication. Now, these peculiarities obviously do not ground any merit and therefore entitle us to no special protection.

Things, however, are different with autonomy. This feature is not only peculiar to, or most developed in, the human species. It also has an intrinsic value, as it shows humans as capable of behavior that exacts respect. We are not merely free;

we are free to act on a principle that we perceive as morally obligatory. And precisely because we have this capacity, precisely because morality is within our reach, we are entitled to an amount of respect unfettered by contingent circumstances.

The argument is largely inspired by Kant and yet it shares with Kant only the intuition of a link between a capacity for moral agency and dignity. Down the argumentative path, whose details I cannot present here, I take a couple of turns that make my approach significantly different from any Kant would be ready to endorse. To begin with, I hold that autonomous agency need not be restricted to agency under the auspices of the Categorical Imperative. I argue that authentic, duty-based agency occurs even when people act under different moral imperatives, such as the Golden Rule or other moral formulas, including the maximization principle dear to utilitarians.[1] Also, and perhaps even more against Kant, I deny that autonomous agency is a peculiarity of human animals, defending the weaker thesis that humans are merely capable of this form of agency *to the highest degree of development* in the animal world.[2]

As it was easy to predict, it is on these two major departures from Kantian orthodoxy that readers have turned their critical eye. In the context of this paper I reply to Luigi Filieri and Natalia Lerussi who both react to my severing autonomy from the Categorical Imperative. Filieri reminds me that the Categorical Imperative is not a law with its own normative content. Rather it is merely the form a law (better, maxim) must have if it must be permissible. So the Categorical Imperative merely prescribes that the maxim you are adopting could be universalized without logical or practical contradiction (formula of universalization) or could be willed to hold as a natural law (formula of the law of nature). In and of itself, however, it prescribes no content or matter. It follows that we do not need to alter Kant's idea that autonomous agency rests on the Categorical Imperative to have all the latitude in the moral law I want to avoid parochialism. If I understand him correctly, Filieri thus wants to question the necessity of liberalizing the boundaries of the moral law's content as to include the Golden rule or other allegedly less parochial formulae for the sake of transcultural validity.[3]

---

1 As it will be shown soon, I have radicalized my position since the publication of the book because I now argue that *any* moral formula can serve as inspiration for authentic autonomous behavior, on the sole condition that it is a general principle that does not merely serve my selfish interests.

2 Obviously for Kant human beings are not the sole entities capable of moral agency. He thinks that purely rational agents or partly rational entities whose existence we may speculate about, like respectively angels or intelligent extraterrestrials, are equally autonomous. Since what confers autonomy is ultimately our rational nature, it follows quite naturally that all rational entities are autonomous. In fact, as Kant puts it: "unless one wants to refuse the concept of morality all truth and reference to some possible object, one cannot deny that its law is so extensive in its significance that it must hold not merely for human beings but for *all rational beings in general* [*alle vernünftige Wesen* überhaupt]" (GMS 4: 408; see also 410n., 412, 426, 431, 442). So when we say that humans are the sole autonomous species, we implicitly mean 'of which we have experience", "part of the sensible world," or the like. This point that humans are autonomous simply because they embody a property (rationality) that other species may share with us will be crucial to show, in the third part, that Kant's position and ours have nothing to do with specisism.

3 For a criticism similar to Filieri's see Klein 2018.

Natalia Lerussi cuts deeper and presses me to determine whether my severing autonomy from the Categorical Imperative does not imply that I am making the law underpinning moral behavior fully arbitrary. In fact, it becomes incumbent on me to define with some precision which moral principles, in addition to the Categorical Imperative, are to be considered as legitimate bases of autonomous agency. I need to define new boundaries for the content of moral principles that, if adopted, are supposed to generate autonomous agency. Most fundamentally, the problem is whether only reasonably well-specified moral principles can inspire autonomous agency or any principle can do that on the sole condition that agents adopt it independently of any possible selfish interest. Is it the case that agents are autonomous even when they adopt heinous 'moral' principles? In the book I gave the example of the sacrifice of Nazi officials that decided not to surrender to the Allies. Is their sacrifice an act as autonomous as to the one imagined by Kant in his famous example of someone who refuses to give false testimony against an innocent man, knowing that this will lead him or her to death (KpV 5: 30)?

Let me reply to my two critics in the order I presented their comments. In response to Filieri, the main point is the following. In construing the Categorical Imperative basically as a formal requirement that any acceptable maxim must satisfy, the impression is that he is understanding it as indifferent to any specific normative content. As Filieri puts it, "the moral law legislates the mere form of an action, whatever its content may be" (Filieri 2018: 591). Obviously this cannot mean that the Categorical Imperative does not discriminate between permissible and impermissible maxims, checking only that they are formally correct, which I trust to mean amenable to universalization or transformation into a law of nature. This would mean that all the CI does is to check that I may be able to live or want to live in a world in which my maxim becomes a universal law. But this cannot be all the CI does, at least not without abundant qualifications.

One way to see this is to focus on what Henry Allison calls 'false positives' (Allison 2011 191–196), that is, examples of clearly and unequivocally impermissible maxims that could be easily either universalized or turned into a universal law. The maxim imagined by Paul Dietrichson and discussed by Barbara Herman (Herman 1993, 113–131) is that of killing babies who prevent our sleep by crying at night. If this were to become a law of nature, it is difficult to detect any logical or a practical contradiction. Let me give another example. Imagine that I want to live in a world in which people deceive one another at every opportunity. A sort of world in which it is universally accepted that if one is smart and clever enough to be able to fool others, he should be allowed to do it. Again, there are defensive strategies a supporter of the law of nature formulation could adopt against these counter examples, but it does appear that no practical or logical contradiction arises the moment we universalize this maxim or transform it into a law of nature.

Although a discussion of this would be way too long, and it would probably intersect the much discussed issue of the equivalence between the three formulas of the Categorical Imperative – an issue that has still excellent interpreters holding opposing views – the point relevant for replying to Filieri is that the CI does contain some 'matter' and not only the sheer form of universalizability. And this matter is best expressed by the formula of humanity, where the 'matter' famously is

the prescription to 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 4:429).[4] Once the matter is not forgotten, the above mentioned false positives immediately vanish. To focus on my own example: it does even if I, or all human beings, want to live in a world in which we cheat and deceive each other any time we have the opportunity to do so. Such a maxim, no matter whether endorsed universally or turned into a law of nature, violates the humanity in ourselves and others. As such it would be disqualified as impermissible. This shows is that the Categorical Imperative does have its specific normative content and cannot be reduced to a merely formal requirement, on pain of making Kant's ethics not only vulnerable to Hegel's famous objection of 'empty formalism,' but more importantly close to a parody. And precisely because CI has a specific content, the question opens up whether we want to make that specific content (respect for humanity as an end in itself) as a necessary condition for autonomous agency.

Natalia Lerussi shares with Filieri an uneasiness related to my liberalized version of autonomy. Yet her worries are more profound.[5] To use her words:

> if autonomy is guaranteed, as Caranti wants it to be, through the determination to act in conformity with any general principle, we are confronted with an undesirable alternative: if the principle in question is not given to the agent by reason, but rather they must decide it voluntarily, the question arises of the *lack of a criterion* for deciding between general moral principles. The consequence is that the decision to subordinate oneself to one general principle or another appears arbitrary. If, on the other hand, the principle in question is given to the agent, not by means of their reason, but rather, for example, through culture, the state or religion, it becomes difficult to maintain an acceptable concept of "autonomy". This second alternative also brings the difficulty that it does not permit us to establish a sharp division between moral determination mediated by, for example, the Golden Rule and by a morally perverse principle, such as the determination to act in accordance with the will of the Führer in Nazi Germany. (Lerussi 2018: 630)

Although Lerussi phrases her criticism in the form of a dilemma, it seems that her basic point is that if autonomy is construed as the capacity to act on any general unselfish principle freely endorsed by the agent, then two consequences (I do not think they are two horns of a dilemma) follow. To begin with, it seems that I am allowing for the possibility that the law come from a source other than our pure reason: tradition, religion, culture, accepted morals and so on. How could one call this autonomy? On the other hand, if agents are autonomous merely because they act on a principle, with the sole condition that it must not respond to selfish interests, then I am opening to all sorts of perverse principles. For example, I am committed to say that a principle of unfettered loyalty to Nazism is just as good Kant's respect for humanity to generate autonomous agency. Both points are taken as self-evidently devastating for my conception of autonomy. The consequence of the first is that my conception of autonomy is stained, to say the least, with heteronomy. The consequence of the second is that I go very close to relativism.

4 References to Kant's works are indicated with standard abbreviations and follow the *Akademie* pagination.

5 For a criticism similar to Lerussi's see dos Santos 2018 and my reply Caranti 2018.

Replying to Lerussi allows me to confess that my position has become much more radical than it was in *Kant's Political Legacy*. While I was well aware of both points in the book, I think I am ready now, in a way that I was not before, to bite both bullets. What interests me is the human capacity to ignore one's selfish interests for the sake of some greater cause, even if this greater cause does not coincide with any of the most popular moral imperatives and the motivation for action is not 'pure' in the sense of originating from the recognition of the sheer obligatoriness of the course of action in question. Just to give an example, I think that the Christian martyrs who refused to worship the emperor to avoid the persecutions carried out by the pre-Constantine Roman authority were sacrificing their lives not because they had the Golden Rule or the Categorical Imperative in mind, but out of a free endorsement to *their* 'greater cause', namely obedience to their God. And I want to say that this conduct is as autonomous as the conduct, to use Kant's example, of the person who refuses to give false testimony against an innocent man just because he thinks this is the right thing to do. Quite in line with common sense, I think that any sacrifice of one's interests for the sake of a greater cause is a potential act of autonomy.[6]

In addition to readiness to sacrifice one's immediate interests, in the book I introduce a second condition for autonomy, namely the ability for this greater cause to pass minimal moral standards. To be frank, I am no longer certain that my theory needs this second condition, which, among other things, brings with it the necessity to have a fairly detailed account of what these 'minimal moral standards' are. In other words, I am now inclined to think that even the refusal by Nazi officials to surrender in Berlin and their sacrifice for loyalty to the Führer should be construed as an autonomous act. I think now that we can legitimately admire this act without compromising in the least our strict and profound reprobation of the intrinsic merits of the cause that inspired it. I also think that this position is closer to our moral intuitions as one may think at first sight. In fact, we do admire acts of 'integrity' and self-denial, independently of the merits of the cause for which people sacrifice themselves. We do admire the 'integrity' of Nazi officials who sacrificed their lives for what they believed in, no matter how much we despise the ideology for which they sacrificed.

Finally, and perhaps less problematically, let me reinforce a point about which I did not change my mind and that is relevant to both Filieri and Lerussi. I think that we should be more 'flexible' than Kant about the standards of moral behavior to hope that our autonomy/dignity based conception of human rights be sufficiently consonant with the diverse moral sensitivities of the world. If we liberalize our conception of autonomy making it equivalent to the individuals' ability to act on principle they recognize as obligatory over and above their immediate or selfish interests, we come interestingly close to a conception of human value that has a chance to enjoy transcultural validity. As I try to show in a dense section of chapter 3 of my book, all major religious traditions, at least if reasonably interpreted, consider individual autonomy in my liberalized understanding as a central component of human worth (Caranti 2017: 95–104). And this ubiquitous presence

6 And I say 'potential' because there should not be, hidden behind the greater cause, some sort of unconscious selfish interest. This possibility was well known to Kant and I concede, as he does, that we can be certain about this.

is no accident. There is, so to speak, an *a priori* reason why no religion can fail to acknowledge the central, if not prevalent, value of individual autonomy. Since adherence to a received set of rules, including divine ones, is worth nothing unless it is free and uncoerced, religions cannot help but pay homage to individual autonomy, whether it does so explicitly or not.

2. Can Fictional Characters 'Prove' Individual Autonomy?

As already seen, I claim that the most promising basis of human dignity is our capacity for moral behavior, that I equate, in agreement with Kant, with our autonomy.[7] Once such a strong emphasis is placed on autonomy, one should be certain that autonomy is within our reach. Unless some compelling reasons are given to believe that we do have this capacity, it will appear as bizarre to let human rights rest on such uncertain possession. In fact, without a solid proof the impression will be that my foundation, in relying on some sort of 'invisible basis', produces more, not less skepticism.

Famously Kant himself was challenged to find a proof of the reality of freedom/autonomy. For some time, as shown by Dieter Henrich (Henrich 1973: 107-110), he thought that he could give a theoretical proof of this property of the will. Then he abandoned this plan as evidently incompatible with the limits he himself had come to set for our cognition through the publication of the first critique. He thus moved in the *Groundwork of the Metaphysics of Morals* (1784) to the second best strategy available, that of arguing that autonomy must be necessarily presupposed as a property of the will of entities who take themselves as agents: "now I assert that every being who cannot act except *under the idea of freedom* is by this alone – from the practical point of view – really free." (Gr 4: 448). In other words, I am free because if I weren't, I should stop considering myself an agent. Later, unsatisfied also by this 1784 strategy, Kant attempted to argue in the *Critique of Practical Reason* (1788), through what most commentators see as a 'great reversal' in the argumentative strategy, that the reality of freedom is to be grounded through an appeal to the 'fact of reason', namely the immediate consciousness we have of the binding force the moral law has on us.[8] Remember here the example of the refusal to give in to the threat by the immoral prince who asks from us a false testimony against an innocent man. On my reading, the fact of reason is precisely the immediate and inescapable consciousness that a) refusing is what would be right to do and b) that we *could* refuse. The immediate consciousness of duty reveals that we *are* free, or better that we *take* ourselves as such.

In short, all Kant was able to do, actually all he could do given the general picture of his philosophy, was to provide an argument 'from the practical point of view' for the reality of autonomy. Far from being a scientific or logical proof, his argument

7 This is what Henry Allison calls the Reciprocity thesis (Allison 1990: 201-213).

8 Although the question is still debated among Kant scholars, the vast majority of them do see the difference between the argument of part 3 of the *Groundwork* and the argument in the second critique, turning on the 'fact of reason,' as a reversal. Of this opinion Karl Ameriks, *Kant's Theory of Mind*, Oxford: Clarendon Press 1982, p.226 and Henry Allison (Kant's Theory of Freedom, CUP 1990, pp. 227-90).

only showed (and was meant to show) that autonomy cannot be taken away from the description of our agency without taking away also the very notion of an agent. Even in the most sophisticated version of the argument, which I take to be that of the second critique, freedom or autonomy is revealed through a fact that cannot be displayed, but at most replicated any time in our consciousness, when we encounter duty and discover that we could follow it in any circumstances. This is not insignificant and yet falls short of something we can literally 'point to' a sceptic.

Bojan Kovačević's extended comments on my book contain a lot of interesting suggestions but I want to take them mainly as speaking to this central problem of the proof of our autonomy (or lack thereof). Kovačević's central idea – I reckon – is that we need to live with the fact that we won't ever be able to obtain a proof of our autonomy. Aiming to that would be an act of intellectual hubris, oblivious of the Aristotelian recommendation to search in a discipline for a degree of precision fit for the discipline itself. In the same way in which it would be absurd to be content with a low degree of certainty and precision in a mathematical proof, it would be equally absurd to ask for mathematical precision in social science, let alone philosophy. Nonetheless, Kovačević continues, this does not mean that we have to take our autonomy as an article of faith. We can collect some hints about the reality of this capacity, that would add up to Kant's proofs 'from the practical point of view', from the work of artistic geniuses such as Leo Tolstoy. Literature at times presents us characters that, despite their being a product of human imagination, can bear witness to the human capacity to make the moral law the supreme principle of their actions.

The Tolstoyian character preferred by Kovačević is the nobleman Nekhlyudov in the novel *Resurrection*. Nekhlyudov leads a life full of pleasures to the extent of becoming bored at every spark of beauty in life and insensitive to human suffering. But when he is called as juror in a case against a prostitute who is accused of murder, and he recognizes the girl he once seduced and abandoned, thereby causing the beginning of all her misfortunes, Nekhlyudov hears distinctly the call of duty and devotes his entire life to mitigate the girl's suffering, up to the point of following her to Siberia and sharing with her the condition of inhumane imprisonment. The obvious moral of the story is that, in giving up all his pleasures and choosing the hardships involved in helping the girl, Nekhlyudov is our autonomous hero.

Of course Kovačević does not take into consideration the possibility that our hero is moved to help the girl by a feeling of empathy or compassion that for Kant would make his helping no less heteronomous than his past pursuing of (boring) pleasures. But this is not a major difficulty. At least on our liberalized account of autonomy, what matters is the capacity to act on a principle to the detriment of our immediate selfish interests. The fact that the adoption of that principle can be somehow linked to some higher and more distant inclination (compassion in this case) does not remove the fact that the agent had to prioritize the feeling of compassion over the inclination to avoid all the pain involved in helping the girl. A moment of deliberation about what should be done, as well as of distance from his immediate impulses was therefore presupposed, something which suffices for construing Nekhlyudov's conduct as autonomous under our liberalized account.

The central point to be discussed about Kovačević's reading is a different one and concerns the extent to which art can provide evidence of our autonomy. Novels

are 'proofs', obviously, only to the extent in which readers can connect to the characters presented and identify with them in a manner not dissimilar to the one Kant expects from us when he presents the above cited 'story' of the noble man who resists the prince's threat. After all, Kant's thought experiment is no less a product of the imagination than characters in a novel. It all depends on whether we find it easy to identify with them and therefore find 'evidence' of the human capacity revealed by their actions.

This opens up the general question of the relevance of examples of moral heroism for the sake of proving our autonomy. Not only literature but also history contains stories of moral heroism. But here comes the problem. Literature and history also have numerous stories of moral abjection. The point is not so much whether the atrocities humans have carried out cancel out their good deeds. This was the way I phrased the problem in the book and solved by insisting that no cruel thing actually done can erase a capacity. Since we recognize worth in human beings for what they are capable of, not for what they have done, this worth cannot be touched by actual human evil. The truly difficult question is rather about a comparative assessment of *capacities*. If we focus on a special *capacity* humans have, shouldn't we also consider humans' *capacity* – as special as autonomy – to reach a level of baseness animals are not capable of?

The answer to this objection must likely take one of the two following forms: either one denies that the bad capacity in question is really special of human beings, at least in the same way or to the same extent in which the good one is. For example, one could question whether cruelty towards members of our species or of others is truly a peculiarity of humans. It would not be difficult to find examples in the animal world of 'cruel' behavior. After all, also cats kill not only for the sake of feeding themselves and many mammals kill or eat their offspring or that of other members of the same species.[9] Alternatively, and more promisingly, one could point out that the goodness of autonomy, its comparative advantage if you want, is not exhausted either by its peculiarity. In addition to that, unlike the capacity for extreme cruelty that always seems to serve some selfish need, autonomy presupposes a distance from our impulses, and a deliberating freedom that may even lead to our sacrifice, that makes it distinct and superior than the other 'special' capacity. Only our ability to follow a moral law severed from our selfish impulses reveals us as demi-gods, finite creators, absolute rulers of our own lives. To use the expression by Pico that Kovačević much appreciates, autonomy and only autonomy reveals a spark of 'divine' in us.

3. Impaired Autonomy

If human rights are grounded on a capacity (autonomy), does that mean that humans who are temporarily or permanently impaired in that capacity do not enjoy the protection of human rights? If that were the case, my foundation would run into a fatal difficulty. Human rights are commonly understood as tools in the hands of the weak to defend themselves from abuses of all sort by the strong. If we end up

9 Here is an incomplete list: hamsters, rats, lions, meerkats, and about forty different species of primates.

denying human rights to people who are particularly weak, like those with mental impairments, then something essential has gone wrong.[10]

Fortunately, neither Kant's moral thought nor my foundation are bound to such a counter-intuitive conclusion. First let's distinguish the cases we are dealing with. In the case of children, one can hardly quarrel with the fact that they are given less rights than adults and that this happens precisely for the fact that their rational capacities are not fully developed. For example, children do not have the right to vote because it is assumed that their ability to think autonomously is not developed sufficiently. At the same time, their potentiality for reaching full autonomous status is part of the reason why they have all other rights (human or not) we usually attribute to people. Actually, sometimes they have certain rights – like access to certain state benefits designed to help their development – adults do not have. Hence we attribute more or less rights to children precisely by using their autonomy (or potential development thereof) as a moral compass.

Analogously, elderly people who have lost in part or full their ability to think are denied certain rights (think of all the restrictions that come with a declaration of *non compos mentis*) and yet keep other rights because we still respect them for what they were once capable to do (think and act autonomously). Respecting a rational creature when its capacity for fully autonomous behavior is in place seems to entail respecting her even when she happens to lose – in part or fully – that capacity. Marc cannot be said to be truly respecting Charles *now*, when Charles is a fully autonomous agent, if it is understood that Marc can do whatever he wants with Charles the moment the latter loses his ability to think. If that is the case, then Marc was not respecting Charles even when Charles was healthy.

The same point can be seen from another angle: Imagine how odd it would sound if I were to tell you: "I respect you because you have this wonderful capacity for moral agency. Hence, I make sure that you enjoy all the rights that come with that status. But also rest assured that the moment that capacity will vanish I will stop considering you a subject of rights up to the point that you are degraded to the level of animals or the like". One can certainly restrict, like in the previous case, the number of rights enjoyed by the impaired individual (we deny the right to vote or to use property to someone after *non compos mentis* is declared). And yet the subject does not lose all its rights. While there is latitude for discussion about precisely which rights (human or not) the person should retain, what matters here is the principle. We cannot ignore her (intact) capacity for suffering, for having interests and needs without affecting negatively, in retrospect, the way in which we were treating her. *Mutatis mutandis*, the same can be said for people who have lost – temporarily or permanently – their thinking ability because of illnesses or accidents of different sorts.

In my mind, a whole different case is that of people who were born with severely diminished rational capacities that we know will not improve in the future or with no

10 Obviously this objection is nothing but the reformulation, in the language of human rights, of a classical complaint against Kant's moral thought. In its general form, the complaint is that Kant links so closely morality and dignity to human reason that he risks denying children, some elderly, individuals with temporal or permanent mental impairment the protection that moral status entails.

rational capacity at all. In these cases, I submit, we are permitted to treat these individuals in the same way in which we treat other sentient animals. And that our respect should be made dependent (and perhaps proportional) to their capacity for suffering.

This line of thought is slightly different than the one advanced by Allen Wood (1998) and Onora O'Neill (1998) to deal with this classical impasse for Kant's moral thought. Wood, for example, argues – and O'Neill agrees – that we should abandon the idea that rational nature is to be respected always as embodied in persons. We should also respect rational nature in general, which entails respecting "fragments of it or necessary conditions of it, even where these are not found in fully rational beings or persons" (Wood 1998, 198). This is what happens with small children and people who have severe mental impairments or diseases which deprive them, either temporarily or permanently, of the capacity to set ends according to reason. My argument is not that we respect the fragments of the rational nature these particular subjects embody. Rather I am arguing that we respect the *full* rational capacity these subjects have the *potentiality* to reach or used to embody. It remains an open question (to me at least) whether Wood and O'Neill would agree with my suggestion that these cases are radically different from the ones posed by human subjects who never were and never will be rational.

Recently, in the context of a reply to Peter Singer's famous attack on speciesism (Singer 1975), Shelly Kagan (Kagan 2016) has made a compelling case to show that one can both identify the source of our worth in rational nature (in my language, autonomy) and be able to attribute to people with diminished rational capacity the same protection we attribute to 'normal' humans. While the argument in my opinion ultimately fails, it contains a crucial intuition that helps us to understand better my defense against the objection under consideration. Kagan starts from the premise that the reason why we attribute to ourselves a higher status than animals is not generically what makes us human, but specifically the part of our genetic setting that constitutes us as *persons*, which Kagan rather loosely defines as a rational capacity and self-awareness. We can see this when we realize that we attribute the same degree of moral consideration to entities that we identify as persons, but that are not human. We do not think, to use Kagan's example, that an evil act against superman or E.T. is less of a problem because these two individuals are not human. The offence is serious because superman and E.T. are *persons* in the specified sense. This intuition, that I fully share, shows why Kantians are not speciesists. Going back to a point we already made, for Kant what confers dignity to us is not our belonging to a species, but the fact that this species, at least in the vast majority of its members, *embodies* a property that non-human entities like angels or intelligent extraterrestrials embody or could embody as well. More importantly, Kagan's emphasis on personhood as the true seat of value helps us to see why impaired autonomy is not an insurmountable obstacle for our (and Kant's) account. The point we made above can now be reformulated as follows: those who have the potentiality to be or used to be persons can never be said to be truly respected unless one extends the privileges that come with that status to the time in which respectively personhood has not fully actualized or is lost. [11]

11 Kagan, however, does not use this intuition to arrive to the natural conclusion regarding humans who were born with no rational capacities. Instead of saying, like us, that these

Conclusion

The relation between human dignity and autonomy, combined with my – we could say – deflationary view of the conditions that make human agency autonomous, is undoubtedly the nerve of the foundation of human rights I offer in *Kant's Political Legacy* (first part). Among the many difficulties my foundation encounters two stand out as particularly acute. One is the problem of what I called elsewhere the possibility of deontological barbarians, that is, people who satisfy my condition of autonomy (they act on principles that do not serve their selfish interests) and yet are clearly perverse from a moral point of view. Nazi officials are unfortunately only one easy example of this human category far more populated than one may think. While in the book I tried to avoid the thesis Nazi officials display autonomy when they sacrifice for the Führer, I am now inclined to accept this conclusion and to see it, perhaps too benevolently, not as a vice, but as a virtue of my account. After all, on any reasonable interpretation of human rights, they have to apply and defend from inhumane and degrading treatment not only good people, but also the worst of us. And focusing on the capacity to sacrifice for a 'higher' cause, no matter how perverse this cause is, may not be the wrong move in order to offer protection even to these individuals.

Secondly, the problem of proving that we are autonomous, obviously not a secondary concern for a position that puts so much emphasis on this faculty, shares all the difficulties originally encountered by Kant. In fact, my liberalized version of autonomy is no less a 'breach' in the natural course of events than Kant's original notion. What I tried to argue, taking myself as an orthodox Kantian here, is that we are obliged to look for a proof 'from the internal viewpoint,' that is, from the way we are not merely accustomed but obliged to look at ourselves. In other words, the proof of autonomy must turn on the self-image of agents that is forced upon us by the immediate consciousness of our capacity to act, and sacrifice if necessary, for a selfless cause. In this vein, fictional characters do contribute to the proof of our autonomy – I think Kovačević's basic intuition is correct – on the sheer condition that our identification with them is strong.

Finally, I replied to a classical objection to Kant's ethics that applies with equal force to my account, although I did not address it in the book. This gave me the opportunity to show not only that the objection is considerably less devastating than what critics think, but also that neither Kant's nor my position are stained by speciesism, a very welcome and unexpected result in the context of a foundation of rights that seem to risk that stain in their very concept.

---

humans are not persons and therefore are not entitled to human rights, Kagan argues that, while these individuals are not 'persons', they are nonetheless members of a 'persons-species', that is, of a species that *most of the times* display in its members the required feature of personhood. In virtue of this belonging to a person-species these members deserve a better treatment than primates or animals in general. Ironically, Kagan seems here to fall in the position he was attacking. If being a member of the species brings to individuals a special value *merely* in virtue of what the species *normally* displays, and that, by definition, these particular individuals born with no rational capacity do not, why should we grant them the same protection we reserve to individuals endowed with the relevant feature? It seems that one could do so only through the endorsement of that speciesism from which Kagan's modal personism was supposed to be sharply distinguished.

Bibliography

Allison, Henry (1990), *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
Allison, Henry (2011), *Kant's Groundwork of the Metaphysics of Morals. A Commentary*. Oxford: Oxford University Press.
Caranti, Luigi (2017), *Kant's Political Legacy. Human Rights, Peace, Progress*. Cardiff: University of Wales Press.
Caranti, Luigi (2018), "Reply to my critics" *Estudios Kantianos* 6 (1) 2018: 117–129.
Dos Santos, Robinson (2018), "Sobre autonomia e dignidade como base para justificação dos direitos humanos" *Estudos Kantianos* (6) 1: 39-48.
Filieri, Luigi (2018), "Autonomy, Dignity and History in Caranti's *Kant's Political Legacy*", *Philosophy and Society*, 29 (4): 586-597.
Henrich, Dieter (1973), "Der Begriff der sittlichen Einsicht und Kants Lehre vom Faktum der Vernunft" in G. Prauss (Ed.) *Kant: zur Deutung seiner Theorie von Erkennen und Handeln*, Köln: Kiepenheuser & Witsch: pp.77-115.
Herman, Barbara (1993), *The Practice of Moral Judgment*. Cambridge MA: Harvard University Press.
Kagan, Shelly (2016), "What's Wrong with Speciesism" *Journal of Applied Philosophy* 33 (1): 1–21.
Klein, Joel Thiago (2018), "Uma Fundamentação Kantiana dos Direitos Humanos, da Paz e do Progresso: Um Debate com Luigi Caranti" *Estudos Kantianos* (6) 1: 97–106.
Kovačević, Bojan (2018), "Subjective Universality of Great Novelists as an Artistic Measure of History's Advance towards Actualising Kant's Vision of Freedom", *Philosophy and Society*, 29 (4): 567-585.
Lerussi, Natalia (2018), "Luigi Caranti, *Kant's Political Legacy. Human Rights, Peace, Progress*", *Philosophy and Society*, 29 (4): 629-633.
O'Neill, Onora (1998), "Kant on Duties Regarding Nonrational Nature: II.", *Proceedings of the Aristotelian Society* Supplementary Volumes, 72: 210–228.
Tasioulas, John. 2015. 'On the Foundations of Human Rights', in R. Cruft, M. S. Liao and M. Renzo (eds), *Philosophical Foundations of Human Rights*, Oxford: Oxford University Press, pp. 45–70.
Singer, Peter (1975), *Animal Liberation*. London: Jonathan Cape.
Wood, Allen (1998), "Kant on Duties Regarding Nonrational Nature: I." *Proceedings of the Aristotelian Society* Supplementary Volumes, 72: 189–210.

Luiđi Karanti

Kant's Political Legacy. Human Rights, Peace, Progress
Odgovori: autonomija i ljudsko dostojanstvo. Preispitivanje knjige Kantovo političko nasleđe: ljudska prava, mir i progres

Apstrakt

Članak se usredsređuje na neke problematične teze moje knjige *Kantovo političko nasleđe. Ljudska prava, mir, progres* (UWP 2017). Preispitivanje je jednim delom podstaknuto komentarima koje sam dobio a drugim delom procesom samo-kritike. Osnovni problem koji su komentatori uglavnom kritikovali jeste moj pokušaj da dovedem u vezu ljudsko dostojanstvo i našu sposobnost za moralno ponašanje, ili autonomiju. Prvi deo članka podseća na osnovna obeležja mog Kantom inspirisanog a ipak u mnogo čemu anti-Kantovog shvatanja veze između dostojanstva i autonomije i odgovara na neke od kritika koje upućuju ortodoksni Kantovci. Drugi deo je striktno povezan sa prvim budući da razmatra razloge zbog kojih mi verujemo da smo autonomni. Dok sam u knjizi skicirao dokaze samog Kanta za "stvarnost slobode" kako on to kaže, sada se usredsređujem na predlog Bojana Kovačevića da se indirektni dokazi u korist autonomije potraže u likovima iz romana kao delima umetničkih genija

(posebno Lava Tolstoja). To mi omogućava da razmišljam od tome kakva se uopšte vrsta dokaza može legitimno očekivati u ovoj stvari. Treće, odgovaram na klasičnu primedbu, koja se u knjizi ignoriše, a koja podjednako pogađa Kantovu etiku i moju sopstvenu poziciju. Problem se odnosi na ljude sa trenutnim ili trajnim poremećajem racionalnih sposobnosti. Ako dozvolim da ljudsko dostojanstvo zavisi od naše sposobnosti za autonomno delovanje, da li se onda suočavam sa kontraintuitivnim (i prilično razornim) zaključkom da deca ili ljudi koji pate od trenutnog ili nepovratnog gubitka racionalnog kapaciteta, i *a fortiori* autonomije, nemaju dostojanstvo i stoga ne zaslužuju da budu zaštićeni ljudskim pravima?

Ključne reči: moralnost, autonomija, dostojanstvo, ljudska prava, Kant, mentalno oštećenje, vrstovnost