

THE DIGNITY OF PERSONS:
KANTIAN ETHICS AND UTILITARIANISM

Lucas Sierra Vélez

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2020

Full metadata for this item is available in
St Andrews Research Repository
at:

<http://research-repository.st-andrews.ac.uk/>

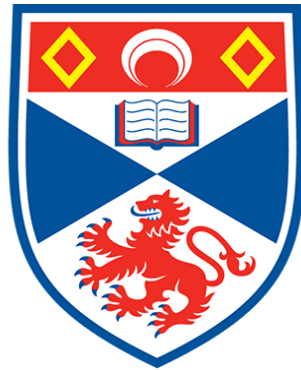
Please use this identifier to cite or link to this item:

<http://hdl.handle.net/10023/20205>

This item is protected by original copyright

THE DIGNITY OF PERSONS:
KANTIAN ETHICS AND UTILITARIANISM

Lucas Sierra Vélez



University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of
Doctor of Philosophy (PhD) at the University of St Andrews.

September 2019

Candidate's declaration

I, Lucas Sierra Vélez, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 76,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in September 2015.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

11 June 2020

Date

Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

9 June 2020

Date

Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Lucas Sierra Vélez, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

11 June 2020

Date

Signature of candidate

9 June 2020

Date

Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Lucas Sierra Vélez, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

11 June 2020

Date

Signature of candidate

Abstract

This work is an attempt to develop a general ethical framework, the product of the synthesis of Kantian ethics and Utilitarianism, giving us a basic account of the dignity of persons and the structure of the moral community. There is a Kantian doctrine of human dignity, and an associated conception of the nature of morality, which together should be taken to provide the basic conceptual framework in ethical theory. However, the conception of morality implicated in the Kantian story recognises not only the dignity of human beings as moral subjects, but also the dignity (the special moral standing) of all sentient beings, who therefore count as persons for moral purposes. This is shown to be the case by means of an argument for Utilitarianism (taken as a technical view in philosophy) being consistent with the basic Kantian framework already in place, argument that involves disassociating Utilitarianism from both Consequentialism and Welfarism and construing it as an Agapist doctrine (i.e. as philosophy of practical love or rationally-based benevolence).

Contents

Abstract.....	iv
Acknowledgments	viii
References to Kant's Works	ix
Introduction.....	1
1. <i>The Core Project</i>	1
2. <i>Utilitarianism, the Dignity of Persons and Dominant Narrative in Philosophy</i>	4
3. <i>Utilitarianism, Agapism and Christian Ethics</i>	8
4. <i>Mill, Kant and The Fundamental Principle of Morality</i>	11
5. <i>Mill, Kant and an Even Larger (Historically Engaging) Project in Moral Philosophy</i>	14
6. <i>On the Use of Labels</i>	16
Part I: Human Dignity: A Defence of Kantian Ethics.....	21
1. <i>The Starting Point: the 'Utilitarian Critique' of Human Dignity</i>	22
1.1 Against Speciesism.....	23
1.2 Against Anthropocentric Ideas in General	27
1.3 Singer's First Premise: The Principle of Individualism.....	29
1.4 Singer's Second Premise: The Sentio-Centric Commitment	30
1.5 Assessing Singer's Critique.....	34
2. <i>The General Kantian Notion of Dignity</i>	36
3. <i>Towards a Substantive Conception of Human Dignity</i>	38
4. <i>The Kantian Doctrine of Human Dignity</i>	41
4.1 Humanity as a Capacity for Morality	42
4.2 Human Moral Equality and the Two Stages of Human Dignity:	44

4.3 Morality as a ‘Dignity Term’	47
5. Interlude: On The Agent vs Patient Distinction	49
6. Against ‘Standard Utilitarianism’ (For Kantian Ethics)	56
6.1 Against ‘Consequentialism’ – Defining the View	58
6.2 Against Consequentialism – The Argument.....	62
6.3 Against ‘Welfarism’ (For ‘Kantian Axiology’)	65
6.3.1 The Intrinsic Value of Happiness.....	66
6.3.2 The Moral Side of Things.....	71
6.3.3 On Moral Goodness	75
Part II: Animal Dignity: A Defence of Utilitarianism	84
1. Kant and the Rejection of ‘Animal Dignity’ and ‘Respect for Animals’	86
1.1 Appraisal Respect	89
1.2 Towards Recognition Respect for All Sentient Beings.....	95
2. The Agapist Project in ‘Animal Ethics’	97
3. The Notion of ‘Duties To Animals’	103
4. ‘Kant’s Dictum’, Practical Love for Animals, and Self-Givingness	113
4.1 Kant’s Dictum	114
4.2 Kant and the Notion of Self-Givingness.....	119
5. Respect for Animals	122
5.1 The Contractualist Challenge.....	123
5.2 A Contractualist Reply to the Contractualist Challenge	125
5.3 Animal Rights.....	128
5.4 Animals (Human and Non-Human) and Their Place in Creation	130
5.5 Summary and Conclusion	131
6. Animals, Babies, Persons, and Human Beings	134
6.1 Animals, Babies, and Persons	134
6.2 Passive Citizens.....	138
6.3 Respect for Moral Personality	140

6.4 Animals equals Persons	142
7. Utilitarianism & the Vindication of Kantian Ethics	148
7.1 Gathering Historical Pieces	148
7.2 On Utilitarianism (and Kantian Ethics): The Technical Formulation	153
Part III (Conclusion)	160
Bibliography	169

Acknowledgments

To Jens Timmermann, my first supervisor, immense gratitude for his commitment and involvement in this project. For his careful reading of my work and his challenging and thoughtful feedback. Also for the many opportunities granted me during my years in St Andrews.

To Tim Mulgan, my second supervisor, my immense gratitude for his commitment and involvement in this project. For a great deal of thorough feedback on the many versions of this work, and for the always wise and useful advice.

Also for their especial involvement in this project, I want to thank Nico Müller, Tom Mckenna, Stefano Lo Re, Jacob Lautrup Kristensen, Pärttyli Rinne, Kristina Kersa, Niel Conradie, Caroline Touborg, Ricardo Miguel, Joe Slater, Elizabeth Ashford, John Callanan, Kent Hurtig, Ben Sachs, Sarah Broadie, Quentin Pharr, Janis Schaab, Bas Tönissen, Michael Walschots and Ravi Thakral.

This work would not be here but for the support of Colfuturo, which partially funded my studies.

References to Kant's Works

I will use the following abbreviations when referencing Kant's works:

<i>Anth</i>	<i>Anthropology from a Pragmatic Point of View</i>
<i>CBHH</i>	<i>Conjectural Beginnings of Human History</i>
<i>CPR</i>	<i>Critique of Pure Reason</i>
<i>CPrR</i>	<i>Critique of Practical Reason</i>
<i>G</i>	<i>Groundwork of the Metaphysics of Morals</i>
<i>LE</i>	<i>Lectures on Ethics</i>
<i>MM</i>	<i>The Metaphysics of Morals</i>
<i>Rel</i>	<i>Religion Within the Boundaries of Mere Reason</i>

Introduction

1. The Core Project

There is more than one thing that this work constitutes. If I am pressed to pick one as the main one, however, my answer has to be the following: this work is a *restatement of Utilitarianism* as an ethical theory. This work is also an extended argument to the effect that all animals – all sentient beings – are *persons with dignity* (persons with a specific type of moral standing and citizenship in our moral community under principles of equality), but as you might expect my defence of this claim is constitutive of my case for Utilitarianism as an ethical theory, a case that consists mostly in trying to *state* the view (which I submit has never been done sufficiently or even adequately).¹

This work also argues for a *Kantian conception of human dignity* and for a distinctively *Kantian conception of morality*, but yet again this is also constitutive of my proposed formulation of Utilitarianism as an ethical theory. Which brings me straight to the nature of this work. My job here is mainly a creative or reconstructive one challenging certain aspects of dominant narrative in philosophy: to restate Utilitarianism as an ethical theory and its distinctive account of the moral standing of all sentient beings, i.e. to explain what these things actually amount to (which may not exactly correspond to what most people presently take them to amount to, and this including most self-identified Utilitarians), by bringing Kant's moral philosophy to bear in positive and significant ways (which perhaps many people would regard as out of place given how much narrative we have received portraying Kantian ethics as the nemesis of Utilitarianism).

In relation to Utilitarianism in particular, this work aims to show that the view is perfectly consistent with Kant's doctrine of human dignity (i.e. properly understood and so stripped from certain claims often though falsely attributed to it, most prominently those affirming the objectification and commodification of non-human animals)² and even with

¹ Thus, again, the task of *restating* Utilitarianism as an ethical theory.

² I submit that claims affirming the objectification and commodification of non-human animals, for instance the claim that non-human animals are 'things' with no value beyond 'price' who are therefore to be treated

his basic account of the nature of morality (a view depicting morality as irreducibly a matter of attitude on the part of human beings themselves, i.e. of things happening in the privacy of human wills and therefore quite independently of felicific consequences and tendencies). Now, this project of bringing Kantian philosophy to bear on the formulation of Utilitarianism – what is essentially a positive endeavour in ethical theory – goes together with a negative project, namely that of disassociating Utilitarianism from both *Consequentialism* and *Welfarism*. Openly challenging customary narrative in philosophy, I argue that Utilitarianism neither determines the morality of human activity on the basis of its consequences nor claims that happiness or well-being is ‘the only intrinsic good’; it certainly doesn’t make the morality (the moral value or moral worth) of actions depend on their felicific consequences (their impact on everyone’s well-being, be it in actuality or in expectation).³ Thus, again, the task of explaining what Utilitarianism actually amounts to, which may not correspond exactly to what most people (for some reason or other) presently take it to amount to.

In relation to my proposed defence of the moral standing of non-human animals in particular (which I am presenting as constitutive of my case for Utilitarianism as an ethical theory), the goal is to explain why this can and should be done in relation to ideas

as ‘mere means’, are falsely attributed to the Kantian doctrine of human dignity even if Kant himself portrayed them as essential. Kant has frequently been criticised both for this mistaken attribution and for the claims themselves objectifying and commodifying non-human animals, and deservedly so. However, the punishment for Kant’s mistakes in this area (outrageous as they are, and damaging as they have been) cannot be to write Kantian ethics off, as Kant’s philosophy contains all manner of positive claims, arguments and conceptual frameworks that can be preserved under interpretations of them that don’t entail the objectification and commodification of non-human animals; claims, arguments and conceptual frameworks that can actually be put to work in the project of defending the dignity of non-human animals. Kant’s authority over ‘Kantian ethics’ (which is threatened by these claims) is discussed towards the end of this Introduction.

³ The argument offered in this area is not only that Utilitarianism is falsely (although to some extent understandably) associated with so-called *Consequentialism* and *Welfarism*; the argument is also meant to show that these views are false, thus making dissociation from them not an option but a necessity for Utilitarians. The argument shows – on the basis of thoughts well expressed by Kant – that there is an intrinsic good of the relevant kind other than happiness and that, relatedly, the moral value of actions depends entirely on their relevant *causes*.

- such as those found in Kant and the larger scholastic tradition in philosophy - of human distinctness (human moral uniqueness in particular) and the special moral significance of our humanity (including ideas about the grounds of moral duty, but also about morality requiring us to relate to human beings in general in a special kind of way) under the rubric, once again, of *human dignity*.

And finally in relation to Kantian ethics, the main goal (which I have disclosed sufficiently already) is to disassociate the view from specific aspects of the story about the dignity of persons that is conventionally told on its behalf, a story that objectifies and commodifies non-human animals, *and therefore, in truth, most persons*, deeming them 'things' with no value beyond 'price' who are therefore rightly treated as 'mere means'. Costumery narrative in relation to this issue has been damaging for all the parties involved, and this includes Kantian ethics itself as a view whose credibility has been hurt: how can people trust a view that under presentations of it widely regarded as standard (and for good reason if Kant is to be the ultimate standard in Kantian ethics) deems permissible acts that *in truth* constitute murder, theft and slavery?⁴ A view that in any case (as I have pointed out)

⁴ My special phrasing here is meant to indicate that I don't take Kant (or any of his faithful followers) to deem permissible acts of murder, theft and slavery under those rubrics. Rather, he gives us a view that deems permissible certain acts that *in truth* constitute murder, theft and slavery, excluding their victims from the relevant scope of reflection. To illustrate the kind of thing Kant is committed to saying: (i) *murder* is always the unlawful killing of *someone* (i.e. of a person); (ii) but a cow, for instance, is not 'someone' but instead only *something*, she (a word - a personal pronoun - that Kant cannot use in consistency with his own principles) is a *thing* not a person; therefore (iii) no killing of a cow (not even an unlawful killing) can possibly count as murder. Now, this point alone doesn't show that Kant's philosophy deems permissible acts that *in truth* constitute murder, nor do similar points show that it allows for theft, slavery, etc. What has been shown (which goes deeper and is perhaps worse) is that non-human animals, i.e. *in truth, most persons*, do not even count (as 'things' that we can use as 'mere means') as potential victims of the relevant crimes in Kant's philosophy. No act in relation to non-human animals as direct victims will ever count as an instance of the relevant crimes according to Kant, therefore, but it remains to be seen (what is a separate issue) whether any such action is in fact deemed permissible by Kant; after all, one could objectify and de-personify non-human animals, and exclude them from the scope of the relevant crimes, on the one hand, and nevertheless forbid anything being taken from them including their liberty and their lives, or forbid all actions that would otherwise - if done unto human beings, for instance - count as instances of the relevant crimes, or in any case - what is actually relevant here - forbid all actions that would *in truth* count as instances of the relevant

objectifies and commodifies the vast majority of persons? Kantian ethics may be dominant in philosophy and enjoy a 'good reputation' even outside the profession, but until its principles be systematically revised (to do justice to the dignity of sentient beings in general) nobody can rightly hold it to be true or even a force for good in the world.⁵

Animal dignity is thus the concept under which the moral standing of sentient beings in general is developed (the standing of all sentient beings, human and non-human alike, as intended beneficiaries of morality), which together with the concept of *specifically human dignity* constitutes the moral general concept (which this work attempts to elucidate, both in its meaning and its moral significance, by means of the integration of Kantian and Utilitarian ethical theory) of *the dignity of persons*.

2. Utilitarianism, the Dignity of Persons and Dominant Narrative in Philosophy

From every Kantian it is expected that she will take the *dignity of persons* very seriously as a foundational concept in moral philosophy, but when it comes to Utilitarians the story is very different: here the expectation is that you should disregard dignity as piece of superstition and persons as of themselves morally insignificant. How come these expectations? I ask. How on Earth can it be generally seen as anomalous for a Utilitarian

crimes, which one can do without recognising them as such). Now, since nowhere in his works did Kant oppose the everyday (by then also customary) consumption of non-human animals, indeed making statements entailing that he had no trouble whatsoever with such consumption (see, for instance, MM 6:409), and given that such consumption is in truth both murderous and theftuous (to mention only a couple of its evil qualities), then we do in fact have grounds to accuse Kant of giving us a view that deems permissible acts that *in truth* constitute murder, theft, etc. Kant is not alone in this, however, as most of our beloved 'classics' (just to speak of them) come out guilty when it comes to doing justice to non-human animals. Not even Bentham (the supposed father of Utilitarianism and according to some an animal liberationist) is off the hook in this area (see, for instance, Bentham 1823/1907, p.311).

⁵ There are timeless truths in Kantian ethics even as Kant gave it to us originally, but these truths also suffer: they suffer from their association (which can of course be undone) with all the doctrine and rhetoric – all the evil doctrine and rhetoric supporting the objectification and commodification of the vast majority of persons – that we rightly blame on Kant's folly. Happily, Kant is not the last word in ethics, not even in Kantian ethics.

to talk about the dignity of persons in relation to the foundations of her own ethical theory? There is a simple answer for this: dominant narrative in philosophy.

Thus, it is very often that we see people criticising Utilitarianism for failing to properly account for – and even for completely ignoring – the *dignity of persons*, but I argue that the truth is the reverse: without Utilitarianism (or in any case without a philosophy that is *duly sentio-centric*, i.e. sentio-centric about the relevant domain of moral significance) it is impossible to provide a credible account of the *dignity of persons*. Without a philosophy that deems the *capacity for happiness* sufficient for personhood, without a philosophy that accounts for the special moral standing of sentient beings in general and the basic moral equality that obtains between them (their equality as *intended beneficiaries of morality*, this being the domain of moral significance where a sentio-centric approach in moral philosophy is mandatory and essential),⁶ what we are likely left with is the objectification and commodification (in philosophy) of the vast majority of persons. The fact that Utilitarians haven't for the most part been using the term 'the dignity of persons' is inconsequential and doesn't in the least undermine the point I am making here, as (i) it is evident that they have been employing the *concept* under alternative headings (e.g. the moral standing of individuals), and (ii) their shunning specific terminology can be explained with reference to sociological factors that have no bearing on positive ethical theorising (e.g. by facts about how terminology has been monopolised in philosophy and related spheres of culture).⁷

We also very often see people criticising Utilitarianism for depicting human beings as nothing more than sensitive beings (for lowering human beings to the level of swine as

⁶ By which I don't mean to suggest that it is mandatory and essential that we adopt the perspective of our sentient (animal) self in our moral thinking, as if the perspective of the sentient self should somehow pass for a moral perspective. All I mean is that the *extension* of the relevant domain of moral significance (namely the domain of *intended beneficiaries of morality* in particular) is fixed by sentience (here as a shorthand for the capacity for happiness), meaning that *all and only beings who are sentient* (i.e. *all and only beings who are capable of happiness*) count as members of that domain.

⁷ Which is not to excuse mistakes that Utilitarians have made in this area, including the very mistake of letting evil doctrine and rhetoric monopolise and define for us our basic everyday terms (the term *person* being perhaps the most important one).

some people enjoyed putting it) and so for denying the rational moral nature (and therefore the very humanity) of human beings. Given that in the final account *humanity equals moral nature* (at least for the purposes of our present discussion), however, the accusation has to be that Utilitarianism denies the existence of human (morally rational) beings, which then raises the question why Utilitarians have been talking about things like *moral duty*. Whose are those moral duties that Utilitarians have been talking about? *Human beings*, of course, whose *moral nature* (and this is the moral fact that is likely obscured when Utilitarianism – or similarly sentio-centric philosophy – is depicted as denying the existence of moral natures themselves, and therefore as having nothing significant to say about the nature of such moral natures) *puts them in relation to sentient beings in general as their intended beneficiaries*, in what I argue deserves to be called an *act of self-givingness*.⁸

Sometimes criticism of Utilitarianism's take on moral natures doesn't get to such wild extremes: maybe the problem is not that Utilitarianism denies the existence of moral natures, of human (morally rational) beings, but rather that it fails to honour these natures in some way, be it by depicting them in poor light (e.g. by failing to emphasise their active side, or by denying their autonomy, or some such), by unduly instrumentalising their importance, or by failing to require action embodying proper recognition of them. The hope is that the argument of this work will give us the bases for addressing these and similar allegations, and the following reflection should suffice for the moment: if in all my moral thoughts I assume my moral nature as something that I cannot but take very seriously, because taking my moral nature seriously is taking my moral duties and my moral mission seriously, what is it that prevents me as a Utilitarian from properly recognising the moral nature of other human beings? What is it that prevents me from thinking that I ought to relate to my fellow human beings in recognition of their moral nature and so not merely as intended beneficiaries of my activity (e.g. also always at the same time as co-workers in the moral community)? Nothing, apart of course from dominant narrative in philosophy.

⁸ The *act of self-givingness* that I refer to here is the success case of the more basic *call of self-givingness* that addresses all human beings: the call (the moral call) to *be kind* and to thereby *act for the benefit of all sentient beings*.

And then, of course, we have philosophers often telling the following story: that while Kantians are the guardians and protectors of absolute prohibitions in morality, and therefore of basic decency, Utilitarians are out to wreak havoc by arguing that ‘anything goes’ (including things like *murder* and *theft*) so long as the ‘felicific calculus’ (perhaps done by some public official, as some people have enjoyed portraying it) dictates. But then, of course, if you think about it for a second (and focusing only on one of the problematic elements in this common narrative), it is plainly incoherent for anyone to allow exceptions to prohibitions against *the unjustifiable*, against things like murder and theft that have *moral unlawfulness* (moral unjustifiability) written into their very definition; and what we end up with as a result of this, at least very often, is Kantians (and others who take themselves to be arguing on similar lines, and therefore against Utilitarians) self-aggrandising, and demonising Utilitarians, while being unable even to materially define things like murder and theft and letting most instances of those moral crimes go under the radar, e.g. in cases where non-human animals (i.e. in truth, most persons) are deemed ‘things’ with no value beyond ‘price’ whom we can do ‘as we wish’ or in any case ‘treat as mere means’, thus placing them beyond the scope of reflection as beings (‘things’) who don’t even get to be considered potential victims of the relevant crimes (of things like murder and theft, but we should also add slavery, rape and more).⁹

⁹ It has to be acknowledged that criticism of Utilitarianism in this area is often (at least to some extent) well deserved. After all, we have seen self-identified Utilitarians often saying that specific moral prohibitions (such as those against *murder* and *theft*, although it has to be said that they have tended strongly to avoid such language when making claims in this area, instead talking about ‘killing’ and ‘stealing’) *cannot be taken as absolute* and that violation of them in general can in principle be justified. Even so-called Rule Utilitarians and Two-level Utilitarians are supposed to be committed to the thought that violation of specific moral prohibitions (such as the ones mentioned here) is permissible and even mandatory in ‘extreme circumstances’. I cannot afford engaging in discussion of these issues on the terms that customary narrative in philosophy has established, as I take such terms to be largely misconceived (especially in what relates to the alleged commitments that go with certain labels, and in particular the thought that Utilitarianism is by its nature, or at least by default, antithetical to the idea of absolute moral prohibitions). As I see it, it should be obvious to *everyone* (not only to Kantians and Utilitarians, but to everyone) that there is no room for thinking about consequences (even intended consequences) that could potentially justify things that have *unlawfulness* or *wrongness* written into their very definition, such as is the case with *murder* (unlawfully killing someone)

3. Utilitarianism, Agapism and Christian Ethics

The developments made in this work are meant to help us combating misleading narrative in philosophy of the kind that I have begun to address, but of course it's the philosophical developments themselves that have to be the focus.

Thus, the reason I appear so confident in claiming that Utilitarianism is very often misunderstood and misportrayed in customary philosophical narrative and even in technical theorising, especially in the ways I already indicated (although in many other ways as well, and this sometimes even by self-identified Utilitarians themselves), is that I have a robust positive story to tell about what Utilitarianism amounts to, a story that has a lot to do with Kantian ethics (in ways that I will be explaining in detail) but which I would rather introduce as follows. In the final account, Utilitarianism is to be defined as an *agapist* doctrine, as a philosophy of *practical love* or *rationaly-based benevolence*. Like Kant's view, and like many other views in the history of philosophy, Utilitarianism depicts morality (the moral quality of human activity) as irreducibly a matter of *attitude* (of acting on good, distinctively moral, motives / principles / maxims) on the part of moral agents themselves

and *theft* (which includes the thought of unlawfully taking something). Thus, it is an open question whether, and if so when and on what grounds, there is ever moral justification for *killing* someone, but there is no real question whether *murder* is ever morally permitted: the very act of declaring some instance of killing *murder* essentially involves the act of thinking of it as a morally wrong (and I should add, *immoral*) act. Such is why we use the term *murder* in moral condemnation, and the same goes for the term *murderer*. *Murder* and *murderer* (and also, of course, *theft* and *thief*) are not morally neutral, but instead *morally charged*, terms, i.e. these are terms that carry a certain – in this case negative – moral valence, and there are many other terms like that (e.g. *lying* and *liar*, *cheating* and *cheater*, etc.). Now, since giving a proper treatment of absolute prohibitions in morality should consist mainly in detailed *moralisation* of specific concepts (i.e. in detailed analysis of certain concepts as *morally charged concepts*), and given that there is no room in this work for taking these concepts one by one in systematic fashion, and given also that treatment of at least some of these concepts nevertheless appears scattered throughout this work, I have decided that there won't be a separate section devoted to addressing the issue of absolute prohibitions in (Utilitarian) morality. You should therefore take this footnote (and the introductory passage in the main text) as my basic treatment of absolute prohibitions and expect the rest of this work to offer a very partial contribution to our understanding of them – which is all that could be afforded – through the moralisation of specific terms (*murder* being the one that receives the most attention).

and therefore not as a matter of these agents producing (or even ‘promising to produce’ in the ‘expected value’ sense in relation to their conduct) good felicific outcomes.

There is of course an essential *reference* to felicific outcomes (or perhaps better to *everyone’s happiness* as a morally obligatory end of action) in Utilitarianism construed as an Agapist ethical doctrine, i.e. the very reference that justifies still using the label ‘Utilitarianism’, which it is my job in this work to elucidate. Something should however be clear before even going into the details of that story: positing a certain reference to something (something having an essential reference to something else) is not the same as *reducing* the object that is said to contain the reference (*morality* in this case) to that other something (everyone’s happiness in this case, or even the conduciveness thereto – a.k.a. the utility – of human activity). The reference to universal happiness in Utilitarianism is *not* going to be that ‘morality equals everyone being happy’, and the same goes for utility or the conduciveness to everyone’s happiness of human activity: it is just not true that morality equals utility (however you take it) in Utilitarianism.

As I am describing it, Utilitarianism’s contribution to philosophy is its description of the fundamental principle of morality as a principle of *practical love* or *rationaly-based benevolence* (agape) involving an essential element of *self-givingness*, of acting *for the happiness of all sentient beings* (this being morality’s essential reference to universal happiness as an end, which justifies still using the label ‘Utilitarianism’). Now, even though Utilitarianism puts especial emphasis on self-givingness thus understood, it makes no claims to the effect that self-givingness is the only constitutive element of morality: Utilitarianism (or in any case the view I am presenting here as one that is meant to deserve the label) offers a substantive and complicated picture of the ethical, which it is my job to describe in detail, and to the largest extent possible also defend, in this work. This description will be drawing heavily on Kant, and hence you will see me depicting morality (still under the label ‘Utilitarianism’, although also under other labels) as essentially involving acts of self-legislation and co-legislation highlighting important aspects of the specifically human role in the moral community.

A fact worth noting in relation to the historical background of this work (a work developing Agapist morality) is that the terms *agapism* and *agape*¹⁰ have been used in philosophy before, even in recent philosophy perhaps most famously in the works of William Frankena (1973; 1976) and Joseph Fletcher (1966). However, and this has to be on the record because it matters a great deal, it is in the context of *Judeo-Christian philosophy* (including the authors just mentioned)¹¹ that ‘agape’ (often, although perhaps not entirely accurately, translated as ‘charity’ and understood as an unconditional, self-sacrificial and indiscriminate – impartially directed to everyone – kind of love, which in the tradition has often been described as first directed to God and then to his creatures as the beneficiaries of God’s love, but which in my story is only assumed to be directed at creatures as intended beneficiaries of moral activity, ‘God’ being understood as a metaphor for the *moral law* of human reason as the source of the relevant command and ‘love of God’ a metaphor for a commitment of *obedience to the moral law*, which I also construe as a commitment to *truthfulness in living*, thus dispelling the myth that practical love is potentially in tension with truthfulness) has been, and still is, more widely and thoroughly treated as a fundamental ethical notion denoting the attitude that is at the core of the morally good life, in something like the way that I am arguing it is. In honour to this particular association, I have decided to make treatment of the fundamental principles of ‘Christian ethics’¹² one of the

¹⁰ Terms that are largely neglected from the perspective of dominant ethical theory in our little corner of the world, except, of course, in theology departments and courses of ‘religious ethics’.

¹¹ The reason for singling out Frankena and Fletcher before introducing Judeo-Christian philosophy as a hotbed of agapist morality (i.e. of moral philosophy being done under that label), is that they are to my knowledge the two most famous cases of modern-day agapist ethicists who are recognised more as philosophers than they are as theologians or religion scholars.

¹² Kant’s peculiar take on *Judaism*, denying it not only the status of *religion* but also status of an *ethical* philosophy, instead depicting it as a political movement requiring only external action, and also his parallel exaltation of Christianity as ‘the only pure moral religion’ (*Rel* 6:125-127), all of which I regard as highly contestable (especially in what relates to the status of Judaism as an ethical philosophy, which I would contest myself), makes it odd for me to speak of *Judeo-Christian* ethics as belonging in the set of labels whose association this work aims to establish, i.e. it may seem odd for me to associate ‘Kantian ethics’ and ‘Judaism’ in particular. I will therefore (conceding to Kant distinctions that may be completely unfair, although also in part for ease of expression) keep using the terms ‘Christian ethics’ and ‘Christian Philosophy’. Again, this is

significant goals of this work, to be done mainly on the basis of reference to familiar biblical doctrines and principles especially through the works of Immanuel Kant and John Stuart Mill.¹³

4. Mill, Kant and The Fundamental Principle of Morality

There are many reasons why Kant should be my main direct interlocutor in this work, and so he is. One of those reasons, which suffices on its own, is that I find Kant's work most useful to discuss (both for good and for bad) when it comes to talking about foundational issues in moral philosophy.¹⁴ However, there is another historical figure whom I also have highlight in recognition of his importance from the perspective of this work, and that is John Stuart Mill. Which is mainly for the following reason: among the classic self-identified Utilitarians of the past, it was Mill who did more of an effort to relate Utilitarianism to other traditions in ethics in constructive ways, including both Kantian and Christian ethics but some others as well. Mill didn't do the job quite exhaustively or even accurately in my

not out of personal conviction, and you should feel free to read me as always speaking of *Judeo-Christian* philosophy.

¹³ Acknowledging that there is a plethora of philosophers and theologians who are very much worth discussing - both for good and for bad - in the context of doing Agapist ethics whom I won't be discussing, including some of the authors I have mentioned. Now, the fact that *agape* (practical love) is - at least to my knowledge - most widely and thoroughly treated as a fundamental moral notion in the Christian tradition (especially if we include classic philosophers in this tradition, and not only the 'premoderns' but also people like Leibniz, Kierkegaard, and Kant, the latter being an example of someone who can be understood as ascribing an important *agapist element* to 'Christian ethics' while not necessarily deeming it an agapist doctrine) than it is in any other identifiable tradition (although I suspect there are very thorough treatments of equivalent notions in other ethical and religious traditions), if it is a fact, is not the same as it being the monopoly of (self-identified) 'Christians' or 'theists' of any kind. To speak of Utilitarians in particular (who may or may not identify as 'Christians' or even as 'theists' of any kind), although they haven't been using the word *agape* very much if at all, there is no question that they have the *concept* in mind because they use it very often in ways that claim it as essential to their own philosophy. I see it being employed in many of the statements that Utilitarians past and present have made about the nature of morality, thinking mostly of those affirming universal or impartial benevolence (another valid name for *agape*) as the fundamental ethical attitude.

¹⁴ I am a Kantian because I find Kant extremely useful in this way, no promise of faithful discipleship needed.

view, perhaps creating more problems than he managed to solve, but he left us with some relevant remarks that very much set the stage for the argument of this work: again, remarks to the effect that *Utilitarianism must be understood in positive relation to other views (or perhaps better, to other labels)* including the ones I have been referring to. Let me then give you some of those remarks and tell you what I make of them (this section focusing on comments referencing Kant, and the next section examining other references).

There are some significant remarks about the nature and significance of Kant's moral philosophy in Mill's work which I will be discussing in the body of this work. However, there is one aspect of Mill's interaction with Kant that I won't be revisiting because it has a purely introductory function for the purposes of my argument, and that is Mill's general comments on Kant's statement of the fundamental principle of morality. Hence, we see Mill saying things like the following (Mill, 1861/2003, p.225): "When Kant (as before remarked) propounds as the fundamental principle of morals, 'So act, that thy rule of conduct might be adopted as a law by all rational beings,' he virtually acknowledges that the interest of mankind collectively, or at least of mankind indiscriminately, must be in the mind of the agent when conscientiously deciding on the morality of the act. Otherwise he uses words without a meaning". Which continues: "To give any meaning to Kant's principle, the sense put upon it must be, that we ought to shape our conduct by a rule which all rational beings might adopt *with benefit to their collective interest.*" What are we to make of these remarks?

Let me first ask. When Mill tells us that Kant "virtually acknowledges" his point about the nature of morality (namely that it contains an essential reference to that the benefit of human beings in general as an end), what is the purpose of the special phrasing in this sentence? One thought is that Mill is trying to *read claims into Kant's principles*, claims that he takes Kant to have made in some vague way or at least gestured towards, or perhaps entailed in some of the claims that he did make (e.g. in the very claim that one's rule of conduct must be fit to be adopted as a law by all rational beings), but which nevertheless he failed to emphasise sufficiently or made sufficiently clear: thus, I suggest, the thought that Kant "virtually acknowledges" the point. Kant "virtually acknowledges" the point in that he came very close to properly acknowledging it.

Another thought (not necessarily inconsistent with the previous one) is that Mill's special phrasing in this context reveals some textual insecurity on his part concerning Kant's actual claims on the relevant subject; which is a shame because Mill had enough textual evidence available to back up his proposed reading of Kant's principles. Expressions like "otherwise he uses words without a meaning", or "to give any meaning to Kant's principle", or in fact "he virtually acknowledges", would all have been needless had Mill been in a position to point to Kant's manifest Utilitarian (or in any case Agapist) moments in, for instance, the *Doctrine of Virtue*. These moments in Kant play a central role in the argument of this work, and I submit they sufficiently establish the presence of an element of *self-givingness* (of acting for the happiness of subjects) in Kant's moral philosophy. Hence, even though Mill didn't bring the relevant textual pieces to bear in support of his proposed reading of Kant, we can still do the job in our own effort to get Kantian ethics right and make positive use of it in our own work (and, if we like, think that we are building on Mill's efforts, engaging in a project that he inaugurated). What matters here is our own ability to bring Kant's moral philosophy to bear in significant ways, saving all that is true and of lasting value in it, in the context of formulating Utilitarianism as an ethical theory.

Mill's interaction with Kant (the part of it that we have seen so far) is not exactly the finest piece of Kantian scholarship, nor would I like to portray it as a good enough guide to for us to revise Kant's philosophy. It certainly has some merit what Mill says in this context, but I find the deficiencies far more notorious. And one of those deficiencies actually comes very close to heart, namely the lack of mention of non-human animals as an important issue in relation to Kant's philosophy, which gives a free pass to what is more deeply wrong in it. It is not so much conceptual poverty on Kant's part that is problematic; it is not for instance that he lacks (because he doesn't) the concept of *intended beneficiaries of morality*, of beings *for whose benefit* moral action (or at least part of it) is characteristically done, which is what Mill seems to be accusing him of lacking *or perhaps failing to emphasise with sufficient force* (which, if put this way, may in the end be a fair point). The main problem (at least from the perspective of this work, although I submit that this is indeed the single gravest mistake in Kant's moral philosophy) is rather that Kant gets it wrong in terms of the *scope of the relevant attitude* (practical love), or what is the same in terms of the *extension of the relevant domain* (the domain of intended beneficiaries of practical love), unduly

excluding non-human animals from that picture. It is here that Kant's philosophy warrants the strongest kind of criticism, and the fact that Mill plays blind is of course very disappointing.

However, again, we don't have to rely on Mill's personal attempt at revising Kant's fundamental principles (as stated in passages that don't necessarily reflect Mill's complete view on the issues anyway, considering other statements that he makes) and are free to come up with our own version of things with the specific focus that we consider pertinent, which in my case has to be targeting the exclusion of non-human animals from the domain of direct moral considerability as the main aspect of Kant's moral philosophy (as stated) which needs amending. I hope it is clear that the purpose in all of this is not to use Mill to correct Kant or the other way around, but rather to draw on both Kant and Mill *at their reconstructed best* (and on a few other philosophers as well) to *help* tell my own story about the nature of morality and the dignity of persons, a story that is fundamentally based on direct thought about the realities themselves (i.e. on my own personal thinking about the nature of morality) and that you should therefore expect to establish its own terms.¹⁵

5. Mill, Kant and an Even Larger (Historically Engaging) Project in Moral Philosophy

Another important thing that Mill does (which I already began to do here) is putting Utilitarianism in close connection with *Christian ethics* understood as an *agapist* philosophy (or in any case a philosophy with a strong agapist component), which he does in such a way as to suggest that Utilitarianism is – as I claim it is – an agapist doctrine along the very same lines: “In the golden rule of Jesus of Nazareth, we read *the complete spirit* of the ethics of utility. To do as one would be done by, and to love one's neighbour as oneself, constitute the ideal perfection of utilitarian morality.”¹⁶ (Mill, 1861/2003, p.194) I thoroughly sympathise with this remark, and in that spirit, and also thinking that the association between Utilitarian and Christian ethics can go in both directions (because why not?), I

¹⁵ I then abide by my basic professional conviction that we ought to do historically engaging, yet always *fundamentally independent*, moral philosophy.

¹⁶ My own emphasis.

will invite you to read this work as giving a basic statement of the fundamental principles of ‘Christian ethics’ insofar as such is understood as an *Agapist* – and therefore in my sense as a *Utilitarian* – moral philosophy (which I leave for others to debate).¹⁷

There is still more to what Mill does by way of relating Utilitarianism to other philosophies (or perhaps more precisely to other labels in philosophy), and I need to examine some of that material here in order to give you a general sense of a larger project in moral philosophy that I suggest we should pursue. Thus, not only does Mill have things to say about Kantian and Christian ethics as we have seen already, going as far as to identify Utilitarianism with the fundamental principle of the latter (and make it consistent with that of the former) as expressed “in the golden rule of Jesus of Nazareth” (Mill, 1861/2003, p.194); he also attributes Utilitarianism to (young) Socrates as represented by Plato in his dialogue *Protagoras* (Mill, 1861/2003, p.181): “[...] when the youth Socrates listened to the old Protagoras, and asserted (if Plato’s dialogue be grounded on a real conversation) the theory of utilitarianism against the popular morality of the so-called sophist.”¹⁸ Perhaps more important to the argument of this work (because here I cannot afford discussing

¹⁷ The association between the two labels here (Utilitarianism and Christian ethics) is of crucial importance for us to sort out the nature of our ethical theories: and not only for the purpose of getting Utilitarianism right, but to do the same with ‘Christian ethics’ as well. However dominant certain quarters within the Christian tradition have been in the context of what is actually a very large (and of course very old) tradition that has many branches, some of which are often underappreciated and even ignored; no matter for how long they have been telling us, for instance, that in Christian ethics only human benefit is considered intrinsically and other animals are reduced to mere means to human ends, which admittedly goes back to antiquity (with an important trend starting in ancient Rome), or that Christian ethics is meant to be the nemesis (yet another archenemy) of Utilitarianism, which has been said over and over ever since the term ‘Utilitarianism’ was introduced, with many people even using the latter as a term of abuse (calling someone ‘utilitarian’ being intended as an insult), the truth is that Christian philosophy has the resources to tell a very different story. A philosophy that claims to take *agape* seriously (an attitude that in truth involves acting for the happiness of subjects quite independently of their moral credentials and nobility) is a philosophy that is by necessity committed to taking the happiness of non-human animals very seriously.

¹⁸ The attribution here seems plausible because there is in fact a superficial Utilitarian flavour to Socrates’ position in *Protagoras*, although I wouldn’t fully commit given the intellectualist overtones of the dialogue and the ever present worry (when it comes many of the ancients, Plato included) of unwarranted eudemonism.

Socrates and Plato) is the fact that Mill also relates Utilitarianism to *Stoicism* and *Epicureanism*, and indeed in ways not unrelated to Kant's interaction with those same traditions in relation to his own philosophy. I should therefore anticipate that the relation between the five labels mentioned in this paragraph (Utilitarianism, Kantian ethics, Christian ethics, Stoicism and Epicureanism) is one of the central themes of this work.

And let's not forget that in *On Liberty* Mill does even some more, here claiming Aristotle as a Utilitarian (Mill, 1859/2003, p.106), as in fact some people rightly or wrongly still do. Mill saw the essence of Utilitarianism everywhere he looked, he thought we need to understand the view in close connection with both ancient and modern traditions in ethical theory, and I am someone who shares that attitude and happens to also think that the way forward for Utilitarians (and for everyone working in moral philosophy) involves giving a prominent place to Kant, by which I mean taking up Kant's philosophy as a major (if not *the* major) historical philosophical resource for the purpose of formulating their views. Let's do Utilitarian theory in historically engaging ways, let's consider *all* the relevant sources and learn from them as much as we can, but let's not miss a single opportunity to do some serious Kantian ethics.

6. On the Use of Labels

This is an openly ecumenical project in moral philosophy, although I also hope one that makes use of labels in meaningful and systematic ways, in ways that can actually help us sorting out the nature of ethics, doing justice both to existing theories and to the things themselves that moral theories in general are supposed to talk about. I therefore need to say more about the use of labels in this work.

It seems to me that the use of labels, and even more so the dialogue between different labels, is a very complicated affair, one reason being the following: while some labels are essentially *historical* in nature (e.g. 'Kantian ethics', which refers to Kant, and 'Christian ethics', which refers to Christ) others are not. Take *Agapism*, for instance, and realise that this label refers directly to a certain attitude (namely *agape* or *practical love*) that is being identified by the theory in question as morally fundamental, without mention of any authoritative figures. And now take *Utilitarianism*, whose implied referent is one of the

most salient defining features of agape as an attitude, namely the act of *self-givingness* (of *acting for the happiness of subjects*, of determining oneself to be *universally useful* with respect to, and *for the happiness of*, all sentient beings) that it involves. Agapism and Utilitarianism are only assuming the moral realities themselves that they refer to and talk about, and hence place no special authority (let alone defining authority) on the hands of historical figures. They are as timeless as philosophical theories can get in that sense. Whether they are true theories is of course a separate issue, but just by their nature they are essentially *non-historical* and can be talked about freely without assuming any kind of reverence for (or even having necessarily to refer to) particular historical figures.

Hence, and this is the concern, how do we propose a conversation between Kantian ethics (what is essentially a historical label) and Utilitarianism (an essentially non-historical one), between a theory that relies heavily on a single historical authority and one that does not? Well, since Utilitarianism has a historical dimension, for there is in fact such a thing as a *movement in philosophy* that people identify by that name (although the details of the story as is often told are all highly disputable), then you can imagine someone suggesting that we historicise Utilitarianism, or in any case focus on what identifiable figures in the movement have had to offer, perhaps giving prominent voice to the ones who coined the term, and somehow compare that with the contributions of the other movement organised around Kant. But of course this project (this duel between sects) would oblige us to first identify the relevant movements themselves and demarcate between them, which appears like a difficult task in and of itself (especially in this case) given that the boundaries seem somewhat artificial and easy to cross. Is the work of R. M. Hare, someone who identified himself as both a Kantian *and* a Utilitarian, part of the Utilitarian or the Kantian movement? Is my work going to be considered part of the Kantian or the Utilitarian movement? Is Kant himself not part of the Utilitarian movement already given his manifestly Utilitarian moments and how often his work has been used constructively by Utilitarians (sometimes even more so than Bentham's)? These questions are hardly of any relevance in the context of positive ethical theorising, and fall well beyond the scope of this work and others of the kind. I see the case for taking the history of philosophy as seriously as possible, but historical philosophical knowledge is altogether methodologically ancillary in the pursuit of positive philosophical truth, *especially in this area*. Others can indulge

talking about movements and sects in philosophy if they want to. In my case, all efforts are devoted to tackling the issues themselves, bringing labels to bear insofar as they are useful in positive theorising (for the sake of fixing certain ideas, for instance), to describe ethical reality as I see it.

In accordance with the above, I have to discard the strategy of historicising Utilitarianism (to look mostly at the history of the movement, if such there is) and then comparing that with the Kantian historical movement, which (I can now say) seems mostly like a *distracting* task: the kind of task we would be given by someone who doesn't want to see serious philosophical progress (e.g. the dismantling of certain kinds of myths and taboos in ethical discourse) being made.¹⁹ My strategy is thus going to be the exact opposite: to 'de-historicise' Kantian ethics, i.e. to take upon myself the job of defining its claims (with the *aid* of Kant, of course, but exercising the right to correct Kant as I see fit on behalf of the theory that deservedly honours his name), and then to put this view in dialogue with similarly 'non-historical' Utilitarianism. To illustrate: I can introduce the basic 'Kantian doctrine of the moral worth of actions', according to which *actions have moral worth if and only if they are done from duty*; now, surely we can think about the nature of this claim and how it relates to Utilitarian philosophy in abstraction from all given narratives about the comes and goes of the history of ethical theory and without even having to rely on the 'primary texts' (which we read once, or twice, or a thousand times, but which can never come to replace our own thinking). Besides, if the Kantian thesis I just introduced for illustration is true, which I submit it is, then it belongs to humanity in general *Utilitarians included*, not to some sectarian movement in philosophy eager to claim it in exclusivity for itself.

Accordingly, then, I will be treating the main labels that are employed in this work (Kantian ethics and Utilitarianism in particular, but also of course *Agapism* as perhaps the central one, and all other labels that are introduced) as ones standing for particular philosophical theses, or perhaps for more complex views and conceptual frameworks,

¹⁹ Not to deny that certain myths are to be addressed on the basis of careful historical work. Myths about the content of ethical theories as developed by particular philosophers (which abound) are a clear instance of this. Historical knowledge in general requires historical work.

which it is *my job* always to define in my own terms. In sum, neither am I treating Kantian ethics and Utilitarianism as ‘traditions’ in moral philosophy, at least not mainly, nor am I simply detaching the views from their existing historical sources. Actually, large portions of my argumentation will consist in meticulous discussion of specific claims and arguments as found in our classical sources, Kant being the central figure in this regard.

7. The Plan of The Argument

As to the structure of the argument, it goes as follows. **Part I** is devoted mainly to establishing a basic Kantian ethical framework that, I argue, is to be assumed in all moral theorising. A doctrine of human dignity is developed which emphasises the place of human beings in the moral universe as the sole moral actors, and then a basic conception of morality is also developed depicting it as irreducibly a matter of *attitude* (of acting from distinctively moral *motives* or *maxims*) on the part of the individual human subject. This is done in dialogue with Utilitarianism as a view that is typically perceived as calling into question ideas of human dignity (and distinctively human equality) and also as depicting morality in a very different way. Consequentialism and Welfarism are here shown to be false, and the promise is that Utilitarianism can be dissociated from them and rebuilt on more solid foundations (namely on the very foundations that are established in this first part of the work).

Part II is devoted mainly to argue for the dignity (special moral standing) of sentient beings in general, which I do in opposition to a negative story in this area that is often told on behalf of ‘Kantian ethics’. None of what is said in this section is meant to count against the credibility of *Kantian ethics* itself; instead, the upshot of the argument in this section is that Kantians are just as obliged as anyone else to acknowledge the citizenship (the beneficiary status) of all sentient beings, and therefore of non-human animals, in our moral community. Besides, the majority of my arguments in relation to the moral standing of non-human animals in this section are based on discussion of Kant’s works (i.e. to the extent that they are based on historical resources at all), and so the argument remains internal to Kantian ethics even in that sense. Even the formulation of *Utilitarianism* itself (as the theory that enjoins the protection of non-human animals in virtue of enjoining the

protection of sentient beings in general) is one where Kant's philosophy is meant to play a major *positive* role. In the end, my conclusion (which should be understood in accordance with the rules of the game as have been established in this Introduction) is that there is no real conflict, but instead *immense synergy*, between Kantian ethics and Utilitarianism as technical views in moral philosophy trying to account for the dignity of persons, the structure of the moral community, the basic facts of human moral relationality.

Part III (which plays the role of **Conclusion** to this work) has a very specific purpose. It adds to the general argument concerning the consistency and complementarity between Kantian ethics and Utilitarianism as technical philosophical views, and therefore also to the 'hybrid theory' that is developed in this work, only now with a focus on more specific issues in ethical theory. The other goal of this section (intertwined with the previous one) is to explain how Kantian ethics and Utilitarianism, and therefore also the synthesis of the two, are meant to relate to *older philosophies* (or perhaps better, to older labels in philosophy).

The work done in **Part III** is meant to do some more damage to the credibility of exaggerated narratives in philosophy portraying Kantian ethics and Utilitarianism as occupying opposite extremes of the ethical spectrum, this time by explaining how the two views (now as *historical* projects or movements in philosophy, demarcation between which remains a difficult task) have a very similar *genesis* in the way they were supposed to *build on past philosophies*. Perhaps the most basic fact in this regard (which I will be discussing in **Parts I and III**) is that Kantian ethics and Utilitarianism were both developed under the maxim of (i) trying to make good of Hellenistic (both Stoic and Epicurean) philosophy and (ii) secularising the basic ethical insights of Christian ethics. My hope is that by the end of the argument we will be able to see more clearly the possibility of a marriage between Kantian ethics and Utilitarianism, the possibility of us being able to transcend a largely senseless and very damaging dispute in moral philosophy.

Part I: Human Dignity: A Defence of Kantian Ethics

Moral philosophy, I submit, is always to be done under two related assumptions: (i) that there is such a thing as *human dignity* (human moral superiority), and (ii) that morality is all about the *realisation* of this dignity (or more precisely about the realisation of *humanity* itself – understood as a mental capacity – as the property which makes human beings morally superior to all other living and non-living beings).²⁰ This is the central claim that I am making on behalf of what I am calling the *Kantian doctrine of human dignity*, a doctrine that is meant to provide the basic ethical framework for the view being developed in this work.

The starting point for the argument in this **Part I** is the apparent *clash* between two camps that we can initially distinguish, i.e. between those who find *human dignity* crucial for morality, and obviously so, on the one hand (the Kantians), and those who think morality better be divorced from ideas of human dignity on the other (the Utilitarians). In light of my general aims in this work as I have been laying them out, I have decided that

²⁰ *Morality* is in fact to be identified with the self-realisation ('the good') of human beings *as human beings* (as opposed to, say, their realisation as sentient beings, in which case what is at stake is *happiness* and not morality, the happy life as opposed to the morally good life). This simple remark helps to anticipate the crux of the argument ahead: the reason I can with such confidence declare my conviction in the moral centrality of human dignity, and therefore of human beings themselves, is that I identify 'the human' (the distinctively human dimension of things, if you will) with 'the moral' (the moral dimension of things), the terms 'human' and 'moral' thus being interchangeable for relevant purposes. Someone's being a human being, for instance, is in this view the same as someone's being a moral (i.e. *morally responsive*, although not necessarily morally good) being. The point, then, is that once the identity between 'the moral' and 'the human' has been stipulated, some notion of the moral centrality of the human becomes inescapable: human beings are exactly those beings who are called for morality, such that we can see morality itself as something *distinctively human*. Now, the main purpose of this work as a whole (my 'agenda' if you will, which relates to the main side theme of this work – namely 'animal ethics' – as well as with my project of restating Utilitarianism as an ethical theory, all of which I am stating upfront) is of course to argue that a proper conception of human dignity should lend no support to notions of human domination or mastery over non-human animals but must instead lead to as radical a view as we can have in this area: a view (a Utilitarian or Agapist view) involving the recognition of the *fundamental equality* – or in other words the *equal dignity* – of all animals (all sentient beings) as *intended beneficiaries of morality* (as beings *for whose benefit* moral activity is characteristically done).

the argument for Kantian human dignity in this **Part I** cannot be developed in isolation from relevant Utilitarian concerns, but must instead interact with them at every point.

Let me then give you (what we could at this point designate) the *Utilitarian critique of human dignity*: the (initial) attempt on the part of self-identified Utilitarians (and other people endorsing relevantly similar sentio-centric commitments) to undermine the credibility of this notion. My version of ‘Kantian human dignity’ will emerge as a response to this alleged ‘Utilitarian critique’, and the central argument of this work will take off from there.

1. The Starting Point: the ‘Utilitarian Critique’ of Human Dignity

The first thing to note is that Utilitarians – in our present context at least – tend to see ideas of human dignity (human moral superiority) as *assuming* religious belief (belief in God) and therefore as suspect. This is a fact about our current state of the art (a sociological fact) which I cannot afford discussing in this work, and one that doesn’t matter a great deal for our purposes given that this work makes no assumptions about the existence of God or his dignity-conferring power (which many people have pointed to as the ground of human dignity). The Kantian conception of human dignity to be developed in this work is a purely secular one that treats God as a regulative idea (as a metaphor mostly), one that can be used for presentation of ideas but which can be dispensed with because the ground of human dignity lies elsewhere (unless, of course, by God we mean the *moral law* as opposed to the moral author of the world, whom actually exists). Given this feature of the view being developed in this work, we can for the moment dismiss the worry that human dignity is a thing only for theists. Morality is not a thing for theists, but for *human beings*, therefore human dignity (the elevation of human beings in virtue of their capacity for morality) is not a thing for theists, but for human beings.²¹

²¹ Of course theists will typically argue that human dignity is something that belongs to human beings *in general* and not just to theists. But the point here is a different one, namely that you don’t have to be a theist (you don’t need to believe in the existence of God) to *coherently believe* in human dignity, since to have that belief you need only believe in *morality* (which clearly you don’t have to be a theist to believe in).

1.1 Against Speciesism

The real, substantive, critique that has been advanced by Utilitarians – the one that is of interest to us in this work – is based on the thought that raising human beings above the rest of creation (‘anthropocentrism’) betrays the kind of *equality in moral standing* that, as they see it, morality essentially involves. Peter Singer’s work on ‘animal ethics’ is one of the most prominent attempts at advancing a critique of this kind, and as far as I know the first systematic one to receive widespread attention in modern academia. I will take up Singer as my main interlocutor in this section, although what I will be discussing is actually a *reconstructed* version of Singer’s views as archetypal of a very common ‘Utilitarian’ (or in any case *sentio-centric*) take on questions of ‘moral standing’ (this with the purpose of adjusting the relevant thoughts to the needs of my argument). It is for this reason that I won’t be quoting Singer directly very much, but mainly explaining his (somewhat reconstructed) views on the issues and discussing what we should make of them.

The first significant idea that emerges in Singer’s critique of human dignity (anthropocentrism) is that advocates of this notion are guilty of promoting what he regards as a *speciesist* attitude discriminating between individuals on what appears to be a morally arbitrary basis: *their membership or otherwise in the ‘human species’*. To be clear, *speciesism* is not to be understood as a prejudice in favour of the members of the human species in particular, nor, therefore, against the members of other (non-human) species. Any kind of *moral favouritism* that discriminates between individuals – be it for or against them – on account of their species membership (whatever that is) is rightly deemed *speciesist* according to Singer. So, for example, the attitude behind the thought that X but not Y can be raised to be killed and eaten, because the former is a pig (a member of *Sus domesticus*) and the latter is a dog (a member of *Canis familiaris*), is meant to be a paradigmatically *speciesist attitude*.²²

²² It seems to me a problematic feature of Singer’s analysis of ‘speciesism’ that he portrays the speciesist as pointing to membership in biological categories as *per se* dignity-conferring; as if people who deny the moral standing of animals on grounds of anthropocentrism, for instance, were necessarily committed to membership or otherwise in ‘the human species’ being that which matters intrinsically about them. I agree that raising someone merely on the basis of species membership *per se* is wrong, and I don’t want to deny that

Now, since we are talking about the credibility of *human* dignity and related ideas, the kind of (alleged) ‘speciesism’ we are concerned with is one that is supposed to be guilty of falsely *elevating human life as such*. Singer’s suggestion, then, is that belief in human dignity (human superiority) constitutes endorsement of an attitude that is rightly dismissed as *speciesism*, and that this is *no different* from the way in which belief in white supremacy constitutes (and is rightly dismissed as) *racism*, and belief in the inferiority of women constitutes (and is rightly dismissed as) *sexism*. Accordingly, he thinks that if we understand what is wrong with *any* of these attitudes, we thereby understand what is wrong with *all* of them.

Take racism, for instance, and the idea that what is wrong with this attitude is that it elevates some individuals above others on the basis of something that lacks moral significance: *their membership or otherwise in one of the ‘human races’*. We don’t regard the ‘white-race’, for instance, as a morally significant category; its boundaries (to the extent that it even makes sense to talk about them) are not morally significant boundaries, and thus we understand that it is indefensible to raise people’s moral standing *because they are ‘white’*. There is nothing morally special about ‘someone’s being white’, is the idea here. Now, you can see how similar thoughts may arise in relation to sexism, and what Singer is suggesting is that the exact same goes for the kind of speciesism that, he thinks, comes with belief in ‘human dignity’. For just like the racist arbitrarily raises the members of a certain race above the members of other races, declaring the former to matter more than the latter, so the speciesist (the ‘anthropocentric speciesist’) goes ahead and arbitrarily puts the members of the human species above the members of all others. And just like the sexist takes *sex membership* to be a morally significant category when in fact it isn’t, so the speciesist takes *species membership* to matter morally in a similar way when in fact it doesn’t. ‘Race’ and ‘sex’ membership, but also in the same manner ‘species’ membership, should all count for *nothing* in our estimation of the moral standing of individuals. It is as simple as that. We

there are people who are guilty of that; however, I also think that very often people who would arguably count as holding speciesist beliefs typically have a *reason* in mind (which of course we assume is a *morally arbitrary* reason the second we call it *speciesist*) posited in connection to species membership *X* as that which does the real moral work (as that which is supposed to explain why *X* is a morally significant category in the first place). I come back to this point in subsequent footnote.

ought to see all attitudes raising individuals on any such basis as equally ungrounded, or, if you prefer, as equally grounded in morally arbitrary patterns of discrimination (Singer, 2002; 2011a; 2011b).

What Singer is doing is quite straightforward and simple. He is pointing out that membership in the human species, no less than membership in a certain sex or race, is such that it is illegitimate to take it into account as a relevant factor when making judgements about the moral standing of individuals. This is a point worth emphasising, for *that* is what Singer takes most advocates of human dignity to be doing: he sees them as raising human beings above the rest of creation (including non-human animals) simply *because they are human beings* – which for him just means: *because they belong in the human biological species*. And thus we get the first argument that we can attribute to Singer in his critique of human dignity: (1) if human dignity being a respectable notion depends on ‘membership in the human species’ being of itself a morally significant category, and (2) if ‘membership in the human species’ is *not* of itself a morally significant category, then it follows that (3) human dignity is *not* a respectable notion. The reason why human dignity is not a respectable notion derives from the reason why ‘membership in the human species’ (someone’s *being a human being*) is not of itself a morally significant category.²³

²³ We have been talking about the speciesist as basing his ‘moral favouritism’ on the facts of species membership *per se*. Under this construal, the ‘anthropocentric speciesist’ would be him who regards human beings as more morally important than other animals *simply because they belong in the human biological species*. However, this kind of *naked speciesism* is by no means the only morally arbitrary way of raising the moral standing of ‘human beings’ above other creatures, nor should we see it as the only form that ‘anthropocentric speciesism’ can take. In fact, it seems to me that most views (most thoroughly ‘anthropocentric’ views) which we might be tempted to call ‘speciesists’ (which here amounts to deeming them *unduly anthropocentric*) don’t single out ‘membership in the human species’ as *per se* dignity-conferring; instead, they typically give us some *reason* posited in connection with ‘membership in the human species’ as the ground of *its* dignity-conferring power, e.g. that all and only human beings (all members of the human species) are image-bearers of God, or that God loves all human beings (all members of the human species) in a way that is different in kind from the way in which he loves, say, pigs and cows. If the reason posited in connection with membership in the human species is not an independently valid reason with some actual bearing in the relevant context of discussion, then the view is rendered morally arbitrary and, I would add, deserving of the label *speciesist*. Views don’t have to appeal to ‘species membership’ *as such* to be deemed ‘speciesist’, is the idea, then. One direct upshot of this reflection is that someone giving us some reason (on top of species membership *per se*) to raise

But why exactly is ‘membership in the human species’ immaterial in this context? You may ask. Singer’s answer is a very simple one: ‘membership in the human species’ is immaterial to the moral standing of individuals because the only thing that can possibly underpin that moral standing is those capacities possession of which, on the part of individuals, would render them proper objects of direct moral concern, which is certainly not fixed by where they fall in a biological taxonomy (as least as we have biological taxonomies defined at present). If you think about it, (i) it is possible for an organism to belong in the ‘human species’ and yet lack any kind of moral standing in virtue of lacking those capacities that are required for it to count as a proper object of direct moral concern (whatever those capacities may be); and (ii) it is also possible for someone to have the required capacities (whatever those are) and yet not belong in the ‘human species’. An unviable ‘human embryo’ is a clear instance of the first case, since – despite being a member of the human species – it lacks all the capacities that philosophers typically single out as dignity-conferring, including sentience (the Utilitarian’s favourite) and ‘moral autonomy’ or ‘moral rationality’ (the Kantian’s favourite). *ALF* (the alien from Melmac) is a clear instance of the second case, having both sentience and moral agency (presumably) despite not being a member of the ‘human species’ (at least if we abide by the way ‘the human species’ is treated in current biological discourse). But if someone like *ALF* (a ‘non-human’ in the dominant biological sense of the word) has a serious claim to moral standing, and an unviable ‘human embryo’ (a member of the human species) has none, then surely

the members of a certain biological species doesn’t get them off the hook of speciesism, since that depends on the moral arbitrariness or otherwise of the proposed reason: if the reason posited in connection with ‘species membership’ is *itself* morally arbitrary, that arbitrariness renders the view ‘speciesist’. There might be other possible (better) ways of talking about these things, but I needed to make this point concerning the nature of ‘speciesism’ (as something we might talk about even if ‘species membership’ *per se* is not really the point of anything) because that is the way we treat similar prejudices: nobody would suggest that offering a morally arbitrary reason on top of ‘race membership’ *per se* to raise the moral standing of the members of a certain race (e.g. that they are typically taller than the members of other races, and that – to make the view complete, as is often done – those among them who aren’t actually taller in this way nevertheless deserve the same privilege because at least they belong to a race whose ‘normal members’ are taller than the ‘normal members’ of other races) gets someone off the hook of *racism*.

membership in the human species is – as Singer would have it – immaterial to the moral standing of individuals.

What of our commitment to human dignity, then? If the respectability of this commitment depends on the dignity-conferring power of membership in the human species being there, and the power is not there, then we have no option but to give it up: failing to account for the moral significance of someone's *being human* (i.e. for Singer, someone's being a member of the 'human species') means that we have to abandon our ideas of human dignity once and for all. Or so, again, the argument goes.

1.2 Against Anthropocentric Ideas in General

Something we could do at this point is take the idea of someone's 'being human' and give it a more philosophical construal having nothing intrinsically to do with membership in *biological* categories, and this move would certainly get us some way (if not all the way) towards deflecting Singer's critique. The thought could just be that some individuals count as 'human beings' (in some technical philosophical sense of the term) on account of a *capacity* of theirs that goes by the name of *humanity*, the clear candidate being the *capacity for morality* that we identify as distinctively human. Thus, advocates of human dignity need only believe that there is a capacity – called 'humanity' – that is *dignity-conferring* on those who have it on account of its moral significance, and no appeals to membership in the human biological species need be made.

But let us not forget that Singer is a self-proclaimed 'Utilitarian', and his critique is motivated by the thought – which seems to stand in opposition to anthropocentric ethical notions in general – that happiness (hedonic happiness) is the *only* thing capacity for which can underpin the 'moral standing' of individuals (the only 'dignity-conferring capacity' to use the language that I am proposing).²⁴ The fact that Singer holds this substantive

²⁴ In line with common philosophical usage, Singer takes the term *moral standing* to mean *moral considerability* or *moral patienthood*. In his view, someone's 'having moral standing' is the same as someone's being 'morally considerable' in a direct kind of way – his being the 'object of duties' if you will – his being someone *to whom* we have moral duties. Accordingly, the thought that a certain property of individuals (e.g. sentience) underpins their moral standing is the thought that possession of the property in question renders individuals

Utilitarian ('sentio-centric' or 'patho-centric') belief means that he is committed to portraying 'humanity' (the property of 'being human', *however it be construed*) as immaterial to the moral standing of individuals: it is because he takes sentient beings *in general* to be morally considerable in a direct kind of way ('to have moral standing'), and because he believes in the related idea of *animal equality*, that he adopts a principled stance against ideas of human superiority in general as they are commonly employed in human dignity discourse.²⁵

Even if we point to *distinctively human capacities* and try to account for their dignity-conferring power, Singer (as a Utilitarian) is committed to taking issue with our proposal: he will insist that no distinctions in 'moral standing' should be made between *sentient beings* with reference to the property (which only some of them have) of 'being human', *even if construed as a capacity for morality*. Let me raise a question to illustrate what Singer may have in mind (acknowledging that he wouldn't put it exactly in these terms): Is it any more murderous to arbitrarily kill an adult human being than it is to take a baby's life without proper justification? The Utilitarian (like myself in this case) proudly says *no*, and explains her answer with reference to the *sentio-centric* or *patho-centric* commitments of her view: both beings have a capacity for happiness (for pleasure and pain), which is all that matters in this

directly morally considerable, which it can do only by making them proper objects (direct objects) of some distinctively moral attitude (e.g. 'respect' or 'practical love'). Now, this way of putting things makes it natural to distinguish between 'entities with moral standing' (beings that are directly morally considerable) and 'entities without' (beings that are not directly morally considerable). Very often I speak this way myself, but that is not to deny other, equally valid, ways of speaking. For instance, instead of '*denying moral standing*' to entities that in our view count as 'things' ('mere means' or 'tools'), we could say that *that* is their moral standing: they have the moral standing of 'things' ('mere means' or 'tools').

²⁵ In the final analysis, Singer's rejection of the capacity for morality ('humanity') as dignity-conferring has to be understood as pertaining to discussions of moral considerability (moral patienthood) in particular; in the spirit of philosophical charity, it should in fact be read as saying something even more specific than that, namely that such a capacity in no way fixes – or contributes to fixing – the domain of *intended beneficiaries of morality* in particular. We certainly shouldn't read Singer as trying to undermine the moral significance of the capacity for morality when it comes to the placing of moral burdens, for instance, or as saying that we ought to relate to human beings in exactly the same way we ought to relate to non-human animals (see Section 5. below). This is not to suggest that Singer does full justice to humanity in his own work, as here we are simply trying to locate the positive contributions of his general critique of anthropocentrism in ethics.

context. And similarly if the question were referring to an adult human being and a dog: it is no more murderous to kill (without proper moral justification, if any there is) the one than it is to kill the other. The question is not: can they reason? Nor: can they talk? But: can they suffer and enjoy their own lives?²⁶

Nothing more needs to be said in order to capture the essence and the spirit of Singer's 'Utilitarian' (or in any case *sentio-centric*) critique of human dignity. Before assessing the merits of this critique, however, and so as to make the ideas a bit easier to handle (and also taking the chance to start sketching a rout for the argument ahead), I need to say more about how this critique is divided into two parts, two separate and distinct claims that are meant to undermine the foundations of 'human dignity'.

1.3 Singer's First Premise: The Principle of Individualism

The moral standing of individuals depends on *their own individual capacities*.²⁷ Such is the first premise in Singer's argument, which we might call *the principle of individualism*. Again, this is the principle according to which we ought to look at each individual separately (in her own individual capacities) to adjudicate on her moral standing. We don't look at the groups (social or biological) where she may belong, but consider only her own existence. For instance, we ask: is *she* a proper object of our benevolence? Which, again, has nothing to do with whether she has certain genes placing her somewhere in a biological taxonomy, or with whether she came from here or from there, or with whether she can reproduce with

²⁶ The point here is simply that for Utilitarians individuals don't gain any 'direct moral considerability' (in the sense of having *beneficiary status* in the moral community) because they are capable of morality. All sentient beings are *equals* in their condition as *intended beneficiaries of morality*, is the idea. Is it any more murderous to kill someone who is a moral actor and not just a person with 'beneficiary status' than it is to kill a 'mere intended beneficiary'? The Utilitarian (like myself in this case) says *no*. Some things may vary from one case to the other, but not the murderous quality of the act: *a person is a person is a person*, and also *murder is murder is murder*, is what the Utilitarian (like myself in this case) is meant to be saying here (see **Part II**).

²⁷ I am speaking of two *premises* here ('two premises in an argument'), but actually I don't mean to be talking about an argument in the proper sense, i.e. one with a few stated premises and a conclusion. What I mean is simply that there are two basic claims on the basis of which Singer builds his general *case* against (what he perceives as) unduly anthropocentric morality (here under the rubric of 'human dignity').

this or that other individual; instead, it is all about whether she is *capable of happiness*, so that we can be benevolent to *her*.

The ‘principle of individualism’ is crucial for Singer because it allows him to argue against the notion of ‘human dignity’ as follows: (1) if the moral standing of individuals depends on their own morally significant capacities, and (2) if there are human beings who entirely lack the relevant capacities (whatever those are), then (3) those human beings in particular who lack the relevant capacities (whatever those are) cannot be granted moral standing. Now, if there are *human beings* who lack moral standing (because they lack dignity-conferring capacities), then ideas of *human dignity* or *human superiority* are straightforwardly false: it is not the case that human beings in general have a moral standing which all non-humans lack.

And so we see how the first premise in Singer’s argument seems to sufficiently undermine human dignity if understood as the idea that all human beings (all members of the human species) have a moral standing which all non-humans (all members of other species) lack. As defined, this notion overlooks the fact that *being a human being* (someone’s being a member of the human species) doesn’t guarantee possession of the morally significant capacities, whatever those are: if it is the capacity for *reasoning and acting morally* that matters, then clearly some human beings (some members of the human species, e.g. foetuses and one-year-olds) lack that; and if it is the capacity for happiness (Singer’s ‘favourite’) that matters, then some human beings (e.g. foetuses) lack that as well, etc. Someone’s being a human being doesn’t guarantee possession of those capacities that suitably underpin the moral standing of individuals by making them proper objects of direct moral concern (i.e. proper objects of distinctively moral attitudes such as, for instance, benevolence or practical love), and this alone should make us suspicious with respect to the notion ‘human dignity’ (at least as defined so far).

1.4 Singer’s Second Premise: The Sentio-Centric Commitment

Singer’s case against ‘human dignity’ has a second premise which derives from his positive moral views, and in particular from his *Utilitarianism* (or in any case from his ‘sentio-centric’ moral commitments). This second premise is the claim that the *capacity for happiness* (often

referred to as *sentience*) is the only property possession of which can suitably ground the dignity (moral standing) of individuals, the only thing that can make of individuals proper objects of direct moral concern. The first premise was a formal claim about the *kind of thing* that can ground an individual's dignity or moral standing, *viz.* his own morally significant capacities as an individual; whereas the second premise – the one we are now examining – is a positive claim about *that* which actually underpins this dignity, *viz.* the individual's own capacity for *happiness*.

I may start to sound too repetitive at this point, but we have to be clear about how the second premise *adds* to Singer's case against human dignity. His first premise tells us that the moral standing of an individual depends entirely on his own morally significant capacities, *whatever those are*; and now we are filling in the gap in the previous sentence by adding the claim that the only dignity-conferring capacity of individuals (the only capacity that can give them some special moral standing) is their *capacity for happiness*. Only those individuals who have this capacity count as having the relevant moral standing; they and they alone are proper objects of direct moral concern. Thus, while Singer's first premise seems to defeat human dignity (anthropocentrism) if understood in a certain narrow way (i.e. as depending on *membership in the human species* being of itself a morally significant category), his second goes beyond that by providing a positive reason to reject notions of human dignity more generally, even if construed in ways that are not easily dismissible as 'speciesist', i.e. even if construed as basing 'human dignity' in the (alleged) moral significance of distinctively human capacities.

Something that follows from Singer's substantive claims about the actual basis of moral standing (his second premise) is a certain understanding of what our notions of *moral equality* (in relation to the treatment of individuals) actually amount to. In particular, he thinks that this equality has to be founded in the *capacity for happiness* which some individual entities (who therefore count as *subjects*) have in common; and once he has emphasised this point, he can point out the fact that this 'moral equality' is one that obtains between individuals belonging to many different species, and therefore not only between human beings. Most members of the human species have a capacity for happiness and therefore participate in the ideas of moral standing and moral equality that Singer (as a Utilitarian) believes in, but obviously they share these features with members of other species insofar

as they *too* have a capacity for happiness. Hence, not only is it the case that not all human beings have moral standing or dignity (as entailed by the first premise in Singer's argument); it is also the case that the moral standing that some human beings *do* have (on account of possessing the relevant capacity) is not one that all non-humans lack, but instead one that a vast number of non-humans have. After all – and here is the punchline – the capacity for happiness is not a characteristically human capacity but rather a *characteristically animal capacity*, one that constitutes the core of what we might call *the animal condition*.²⁸ Now, if one agrees with Singer in thinking that this 'characteristically animal capacity' is the only one to consider when making decisions about the moral standing of individuals, then of course one's suspicions against human dignity as a respectable notion will be nothing but amplified, for now one has the thought that moral standing is something that characteristically belongs to *animals*, not to human beings. Well, of course it belongs to human beings *qua animals* (or in any case *qua capable of happiness*), but that is a different, unproblematic claim for a Utilitarian like Singer: dignity (moral standing / moral patienthood / direct moral considerability / end-in-itselfhood) goes where the animal goes, and the human animal is only *one among the many*.

I began describing Singer's views by considering his claim that advocates of human dignity are guilty of a moral mistake of a terrible kind: in their *speciesism* – Singer claims – they are the equivalent of the sexist who believes that men alone ought to rule, or the racist who believes that white people are to be treated as more important and deserving than the rest. As we have begun to see, however, this applies only to (i) those who make human dignity depend on 'species membership' as understood in biology being a dignity-conferring

²⁸ We don't suffer or enjoy as human beings; we suffer and enjoy *as animals* (i.e. as *sentient our experiential beings*). In suffering and enjoyment we are the equals of cows and pigs, sheep and mice, turtles and lizards. Now, later in the argument (towards the end of **Part II**) I will be questioning the identity – which I am proposing here in the main text – between 'animality' and the 'capacity for happiness' (or sensibility more generally, also including concomitant agency), which I will do on the basis of the thought that 'animal' turns out to be a more general philosophical category, one that is actually equivalent to those of 'person' and 'subject'. To prevent misunderstandings, let me note that (at least for now) whenever I say 'animal' I simply mean 'sentient being', and when I say that something is 'distinctively animal' I mean 'distinctive of sentient life'.

feature of individuals and (ii) those that posit morally arbitrary reasons in connection with species membership to justify their moral predilections; it shouldn't concern those who in a more philosophical attitude think of humanity as a *capacity* (on the part of particular individuals) that is of itself morally significant in the right kind of way. If your view puts the ground of 'human dignity' in the direct moral significance of such a capacity as instantiated by particular individuals, there is no argument against you from Singer's first premise. Now, your view could still be challenged on the basis of Singer's second premise, which expresses what are normally regarded as distinctively Utilitarian (or in any case sentio-centric) commitments that are meant to stand in opposition to *all views* excluding non-human animals from the scope of dignity (direct moral considerability) or even portraying them as inferior to human beings in that regard.

And so what we get so far is an apparent *clash* between two different takes on the same issue, between two contending philosophical views on the actual basis of our ideas of 'dignity' and 'equality': Singer's Utilitarian ('sentio-centric' or 'patho-centric') view and the 'anthropocentric' ('Logo-centric') view that we find in philosophers like Kant. This is the same clash we pointed out at the beginning, between those who take distinctively *human* dignity to be crucial (the Kantians) and those who supposedly think otherwise (the Utilitarians). By now, we have at least learned where Utilitarians are coming from in their critique of human dignity, which is not exactly a place of amoralism. The reason why Utilitarians oppose ideas of human superiority (human uniqueness and distinctively human equality) is simply that they care morally about animals (sentient beings) *in general* and consequentially see nothing special, nothing particularly morally significant, in human life as such that could render human beings more morally considerable than (superior in 'moral standing' to) other animals. When it comes to deciding whether some individual is a proper object of direct moral concern, they simply think there is no merit to someone's *being a human being* (however this notion be construed).

Singer's Utilitarian critique of human dignity has now been stated in sufficient detail, and now we shift focus more towards assessing the merits of this critique. At this point we begin to introduce elements pointing in a Kantian direction.

1.5 Assessing Singer's Critique

My goal here is to offer a distinctively Kantian reply to Singer's Utilitarian critique vindicating the notion of *human dignity*. However, it wouldn't be fair for me to start pointing out the deficiencies, as I perceive them, of this alleged 'Utilitarian' critique (or talking about the riches of Kantianism in this area) without first saying more about the truth that it contains.

The first thing to be said in favour of Singer's critique (which I have done in passing already) has to be that his *first premise* is on target: the moral standing of individuals – we have to agree – depends on the morally significant capacities they have as the individuals they are. Not all philosophers seem to agree on this basic point, but I think this is one of the places where Kantians and Utilitarians have most reason to agree: we have all the material to talk about the dignity of individuals in those capacities for *morality* and *happiness* that we care about, which are always capacities of particular individuals. We can on this basis agree on the imperative to oppose all attempts at raising the moral standing of individuals on account of things other than *their own morally significant capacities*: their possession of mental capacities that would render them proper objects of direct moral concern, i.e. proper objects of *morally required attitudes* such as, for instance, *benevolence* or *practical love* (or some other attitude that we manage to identify as morally necessary).

There is more to be said in favour of Singer's critique, I think. Recall that his second premise includes the claim that *sentient beings* have a moral standing of their own in virtue of their capacity for happiness. I wholeheartedly agree with this claim, and so **Part II** below is meant to offer – among other things – a defence of what I call *animal dignity*: the idea of a certain kind of moral standing (the standing of *intended beneficiaries of morality*) that pertains distinctively to sentient beings. The problem, however, is that the premise *itself* (the second premise) can be challenged by a well-grounded theory of human dignity – like Kant's – accounting for the special moral significance of humanity as a capacity for morality (often called 'freedom').²⁹ Believing that all sentient beings have a certain dignity or moral

²⁹ Later it will become clear that Singer cannot coherently sustain commitment to his second premise *as stated so far*, since everything he says by way of *moral* claims addresses our human (moral) *reason* and so presupposes

standing (whose nature is yet to be explicated) is not the same as approving of Singer's second premise, which asks us to believe that nothing more needs be said about the moral standing of individuals. Thus, the reason I reject this premise is that I don't think the 'capacity for happiness' is *the only* capacity capable of conferring dignity (moral standing) on individuals. If the 'capacity for happiness' is *not* the only dignity-conferring capacity, because there is another capacity that is also dignity-conferring (namely, as I will argue, the *capacity for morality*), then it follows that the moral standing that pertains to sentient beings as such and in general (the standing of *intended beneficiaries of morality*) cannot be the only kind of moral standing there is. Take the positive 'Utilitarian' story in this area ('sentio-centrism' about the domain of intended beneficiaries) as seriously as it deserves, and you should still find plenty of space for a *substantive* conception of human dignity capturing the distinctive moral standing we ascribe to *human beings* (which, initially at least, seems to go unacknowledged in Singer's expressed views on questions of moral standing).³⁰

In sum, the argument for Kantian human dignity has to conform to Singer's principle of individualism, i.e. it has to respect the claim that an individual's moral standing depends entirely on his own morally significant capacities (on her having the capacities that would make her a proper object of morally required attitudes), and it also has to proceed

our *moral uniqueness* and our consequent moral (strictly moral) superiority with respect to the other animals. In Section 5. below (the **Interlude** section), Utilitarians will get a fair chance to clarify their position and qualify the target of their alleged 'critique of human dignity'. To anticipate, the critique is best understood (and under this construal I subscribe to it) as not necessarily targeting the idea of 'human dignity' (human moral superiority) *itself* but rather a certain (very common) *misuse* of it in 'philosophy' and elsewhere: the going from human moral superiority to either (i) the objectification of animal life (with an emphasis on non-human animals) or (ii) the idea that human beings are in any case (radically) *more important* than other animals in their condition as *intended beneficiaries* of morality. It is in relation to the latter domain (the domain of intended beneficiaries of morality, of beings *for whose benefit* morality is characteristically done) that anthropocentrism has no rightful place in moral philosophy.

³⁰ It seems to me that Singer's first premise may be accepted, and *speciesism* as such rejected, and ideas of 'distinctly *animal* equality' defended (thus doing justice to the truth contained in Singer's second premise), without in any way threatening distinctively philosophical conceptions of human dignity of the kind that we find in Kant (abstracted, as it should be, from its potentially speciesist moments and common *misuses* of it towards the objectification – or less than full moral appreciation – of non-human animals and even of animal or sentient life as such).

cautiously not to exclude non-human animals from having dignity (moral standing) altogether, for otherwise the whole point of **Part II** below would be undermined. What is more directly relevant to the argument: I need to explain what is meant by the term *humanity* if not ‘membership in the human biological species’, and so also what is meant by *human dignity* if not ‘the elevation of the members of the human species above the members of other biological species’. I have gestured towards a certain understanding of the term ‘humanity’ (humanity as a *capacity for morality*), and this already gives you a hint of where I am going with the argument. However, I cannot begin to make my case without first justifying the use of the term *dignity*.

Although we have a natural understanding of the term ‘dignity’ (which is why I have taken the liberty to employ it hoping that my reader will understand the meaning of my claims), the truth is that use of it in philosophy has become a problematic affair (due in large part, or so I would argue, to the grave misuses of it in philosophy and elsewhere that we have been witnessing). In light of this fact, we need to have a basic definition of the term ‘dignity’ before proceeding any further.

2. The General Kantian Notion of Dignity

The concept of *human dignity* that I am developing is an application of the general concept of *dignity* as I find it employed by Kant. Now, people say all sorts of things about what Kant means by ‘dignity’, but most of the time what they have in mind is *human dignity*, which (again) amounts to an application of the more general concept I am trying to clarify in this section. Thus, in its naked form, ‘dignity’ can be defined as the *elevation* of an object above another object along a particular dimension, or in some particular respect.³¹ This is a concept that applies to many things of many different kinds, not just to human beings as entities that occupy a special place in the moral universe (*human dignity*).

³¹ This definition is found in Oliver Sensen’s *Kant on Human Dignity* (Sensen, 2011). Sensen identifies Kant’s general concept of dignity as described here (the concept of something being elevated above – or being higher in rank with respect to – something else) with the ‘Stoic’ concept of dignity (Sensen, 2011, p.143).

You might be surprised to hear this, but the Kantian concept of ‘dignity’ (i.e. the concept of ‘dignity’ that Kant employs in his writings) applies to all things that can be said to be raised or elevated (normally *in kind*) above others in ways that are relevant in particular contexts of philosophical interest, and there can be many such relevant contexts. It applies to all things that *occupy a higher station*, or that *are higher in rank*, with respect to others along some philosophically relevant dimension, and there can be many such dimensions and as many relevant rankings and hierarchies. So, for example, Kant claims that *mathematics* has dignity with respect to the empirical sciences in virtue of some idea of its paramount theoretical importance (CPR A464 / B492). Likewise, I’d like to propose the thought that *philosophy* has dignity with respect to other disciplines in that it is responsible for addressing the most vital and important questions in human life (such that, for instance, it is philosophy and not biology or physics which tells us what it means to be a human being).

Please note how the general notion of dignity as defined above is being applied in these cases, taking my claim about the dignity of philosophy as an example: I suggested that a certain object (*viz.* philosophy) is raised or elevated above another object (*viz.* other disciplines) along a particular dimension (*viz.* centrality of the subject matter of various disciplines in human life), and that *in that sense* philosophy has a certain dignity. I don’t care whether or not you agree with this statement about the dignity of philosophy or with Kant’s claim about the dignity of mathematics, since we are just trying to see how the general concept of *dignity* is supposed to be applied.

Now think about the claim – which Kant also makes – that the *moral law* has dignity (MM 6:397). I make sense of this claim by saying that *the moral law* (**first object**) is more important than *other practical principles* (**second object**) in that the commands of the former are to be given *absolute practical priority* over the prescriptions of the latter, where the **dimension of elevation** is *the pre-eminence of various practical principles in terms of their practical priority*. The moral law is thus *raised* or *elevated* above other practical principles in a specifiable and seemingly relevant way, and hence we talk about it as having a certain ‘dignity’. I take it that these claims about the ‘dignity’ of various objects are easy to understand and not at all mysterious (or superstitious), even if it is possible to disagree with them. *In itself*, therefore, the Kantian concept of *dignity* is completely unmysterious and very easy to apply: you just need two objects (or two classes of objects) and a certain dimension

along which one of these objects is, by virtue of some relevant feature of it, raised or elevated in rank above the other.

3. Towards a Substantive Conception of Human Dignity

Before examining the Kantian conception of human dignity, let us appreciate how Utilitarians like Bentham and Singer could make use of the *general* Kantian concept of dignity (as just characterised) in order to craft their own formulation of *human dignity*. They could argue – as in fact we have seen them doing already – that *human beings* (**first object**) have dignity in the sense that, *as sentient creatures*, they are raised above *non-sentient entities* (**second object**) in terms of *moral standing* (**dimension of elevation**). Surely all sentient human beings are (from the moral point of view as typically depicted in Utilitarian philosophy) *radically more important* than all those entities in the world who are incapable of suffering and enjoying their lives, and this, I think, should be enough to settle the issue whether there is room in Utilitarianism for notions of ‘human dignity’. There is a property to human beings (*viz.* their capacity for happiness, to the extent that they actually have it) that makes them radically more important than – radically higher in moral standing with respect to – all those beings in the world which lack that property. In the final analysis, the moral significance of this capacity that Utilitarians single out as paramount lies in the fact that possession of it can alone render individuals proper objects of the kind of *benevolence* that morality essentially involves (which following Kant I call ‘practical love’).

These ideas will be further developed in due course, and for now I’m just pointing to them as plausibly offering a ‘distinctively Utilitarian’ conception of *human dignity* – or, if you wish, to be more precise, the *dignity of human beings*. Human beings are among the creatures *for whose benefit* – as Utilitarians see it – moral action is done, and this sets them apart – in what looks to be a morally significant sense – from all entities that don’t belong in that domain.

Following this line of thought, it seems evident that Utilitarians can draw a sharp distinction between ‘persons’ and ‘things’, between ‘ends in themselves’ and ‘mere means’, or in any case between ‘intended beneficiaries’ (individuals *for whose benefit* things are morally to be done) and ‘resources’ (objects that are to be used *for the benefit of individuals*).

Should Utilitarians be precluded from using terms like *dignity* and *persons* in this context? Honestly, I don't see why they should: having done the job of precisely defining the term *dignity* in particular (which in the relevant context is necessarily connected with the term *person*), it should be evident to everyone that it belongs to Utilitarians as much it belongs to anyone else.

We can of course expect 'orthodox Kantians' (if I may use this expression) and others who take themselves to be arguing on similar lines in this area, which very often goes with arguing against Utilitarianism (e.g. most 'conservative Christians', if I may also use this expression) to protest, as they have been monopolising over human dignity discourse and may want things to continue the same way. Nevertheless, I submit that they have no right to this monopoly. Utilitarians are entitled to make use of whatever terminology ('Kantian' or otherwise) allows them to make sense of their deepest ethical commitments, and I have begun to speak on their behalf by introducing a certain conception of *human dignity* (or, if you wish, to be more precise, 'the dignity of human beings') that is part and parcel of Utilitarian philosophy as we know it already: the idea that human beings have 'dignity' (elevation in moral standing) *as sentient creatures*, on account of the *capacity for happiness* that makes them proper objects of the kind of benevolence that morality requires at a fundamental level. Surely human beings are not just 'things' or 'tools' or 'resources' in Utilitarian philosophy: they are *intended beneficiaries of morality*.

Now, it is true that in this Utilitarian picture the 'dignity of human beings' is not something that belongs to them in exclusivity, but rather something that they have in common with other sentient beings. However, it doesn't follow from this fact that *human dignity* is thereby eliminated or denied, since one of the things the words 'human dignity' can mean is *the dignity of human beings*: the dignity that human beings have on account of the totality of their being, in virtue of whatever dignity-conferring capacities they may have as individuals. Is 'the dignity of women' (i.e. the dignity of women *as persons*) undermined by the fact that they have it in common with men, or maybe by the fact that it is not grounded in their 'womanhood'? Of course not, and similarly in the case at hand: the dignity of human beings (as Utilitarians understand it) is no less of a dignity because they have it in common with other animals (other sentient beings) or because it is not grounded in their humanity. The human being's 'animality' (his sentience or - what is the same here

– his capacity for happiness) is no less part of his being than his humanity is, and so it can in principle be true that *his* dignity (or a part of his dignity) attaches to him on account of *his* animality (his capacity for happiness in particular). Let me emphasise that in this view human beings are not the only creatures we have to dignify: if human beings have a certain dignity *as sentient creatures* (as beings capable of happiness), then we have to acknowledge that this dignity pertains to *all sentient creatures* (i.e. to all beings capable of happiness).

Now, if that which we have in mind when we speak of *human dignity* is something that human beings have in virtue of their *humanity* (on account of that which makes them *human*), which is supposed to raise them even above other animals in terms of moral standing, then it is not clear that the Utilitarian view can give us what we want. For even if Utilitarianism can make some sort of sense of the notion of ‘human dignity’ with reference to the capacity for happiness that defines the ‘animal condition’ in which human beings participate, we have grounds to complain that something crucial has been left out. Something crucial is missing in the picture when human dignity is *reduced* to the purported ‘*animal dignity* of human beings’, which is what thinkers like Bentham and Singer seem (at least initially) to be doing. To their credit, animal dignity (and so the ‘animal dignity of the human being’) will prove to be a real and very important thing, but that is not the kind of dignity we are concerned with in this **Part I**.

My claim here is that we should endorse a *substantive* (non-reductive) conception of human dignity according to which human beings are morally more important than non-human beings (and radically so, although in a way that has to be specified very clearly and that has nothing to do with their happiness counting for more for purposes of our being benevolent to subjects in general) on account of their *humanity*, i.e. on account of that which makes them *human* (their *distinctively human capacities*). This substantive conception of human dignity I take to be Kant’s, or at least sufficiently Kantian in spirit to deserve the labelling.

4. The Kantian Doctrine of Human Dignity

We have been applying the Kantian concept of ‘dignity’ in various claims concerning the moral standing of individuals, and one idea that has emerged is that of *animal dignity*:³² the idea of a certain dignity (a special kind of moral standing) that, it is said, pertains to individuals on account of their ‘animal’ or ‘sensitive’ nature (in their condition as beings *capable of happiness* in particular). This notion brings human beings under its scope insofar as they have the relevant capacity, and hence we seem to be justified in speaking of the ‘animal dignity of human beings’. This is not the kind of idea that I am pursuing in this **Part I**, however. Instead, my goal is to argue for the existence of a *distinctively human* kind of dignity: the idea that human beings (first object) are – merely on account of their *humanity* – raised above *non-human beings* (second object) in terms of *moral standing* (dimension of elevation).

As you can see, it is a certain idea of *human moral superiority* that is at stake here; although, I repeat, we are yet to see any clue about what this superiority amounts to, what it entails and what it doesn’t entail, what follows from it and what doesn’t actually follow. A good way to start describing the fact of human moral superiority, to which corresponds the thought that human beings have a special kind of moral standing, is to introduce the notion of a *special position* or *station* in the moral universe that human beings alone occupy, and a special and distinctive *role* associated with that position which human beings alone are called to play in virtue of their *capacity for morality*. After all, human beings are exactly those beings who are called for morality, meaning of course that all properly moral roles in the moral community are by definition human roles (roles played by human beings).

³² To clarify again, this is the *dignity of sentient beings as such*. One of the reasons I have decided to stay with ‘animal dignity’ as the main designation is that it allows me to speak, in a way not so tedious, about things like ‘the *animal dignity* of human beings’ and ‘the *animal rights* of human beings’. For my final treatment of ‘animality’ as a morally significant category and even as a general philosophical category, see 6.4 in **Part II**.

4.1 Humanity as a Capacity for Morality

What is it that gives human beings their special dignity and raises them morally speaking above the non-human world? Kant's answer to this question is very simple, namely: their *humanity*, and hence his *substantive* (non-reductive) conception of human dignity. Human beings have a special kind of dignity merely in virtue of their *being human*. But what does Kant mean by the word *humanity*, then, which we have established cannot just be membership in a certain biological species? What does it mean to *be human* in the Kantian picture?

Any answer we give to the latter question will be disputed by some interpreter of Kant or other, but in my mind it is clear what the answer has to be (in this context, at least).³³ In the *Groundwork of the Metaphysics of Morals*, Kant equates humanity with *rational nature* (G 4:429); and by saying that human beings are in essence rational creatures (in the sense that is of relevance in this context) he just means that they have a will that is responsive to the *moral law*, a will that indeed is to be regarded as the *legislator* of such a law (the authority behind it) and is therefore *autonomous* or *self-legislating*. Thus, humanity and

³³ I don't mean to suggest it is a simple task to talk about the meaning of the term 'humanity' in Kant's works. In *Religion within the Boundaries of Mere Reason*, for instance, Kant treats 'humanity' as a non-moral kind of rational condition and proposes that we use "personality" for the moral kind (Rel 6:26–28). I cannot properly discuss the merits (or even the point) of this conceptual move in Kant. I will only say that it is immensely problematic from the perspective of my own project (given how I define central concepts) because it is part of my project to argue that *all animals (all sentient beings) are persons*. In fact, it is my view that no animal is less paradigmatically a 'person' than any other animal (e.g. Babe the pig versus Socrates), and this for two reasons: (i) because 'sentience' (animality as I have been treating it so far) is sufficient for personhood, and (ii) because in my final analysis 'animal' and 'person' are actually to be identified as the same general philosophical category (which of course calls into question the identity between 'animality' and 'sensitivity' as I have been treating it thus far – see 6.4 in **Part II**). Now, Kant could have – in *Religion* – introduced the categories of 'animality', 'humanity', and '*moral* personality' as all belonging to 'personhood' or 'personality'. Such a taxonomy is one that I could potentially endorse; what I cannot accept is reserving 'personality' (at least if taken as synonym for 'personhood' or 'someone's being a person') for human beings on the basis that they alone are moral beings. Thus, again, my decision to stay with 'humanity' as treated by Kant in *Groundwork*, which of course you can equate – as I am doing right now – with '*moral* personality'. This decision makes sense of our reference to *human dignity*.

a *capacity for morality* (i.e. for self-legislated moral action) are one and the same thing as Kant understands these notions: that which makes us *human* – in the technical sense of the term that is of relevance in this context – is our *responsiveness to the moral law*, which allows us to become *morally good* or (what amounts to the same thing) to exemplify *moral goodness* in our conduct.³⁴

Here we could also speak of the human being's *freedom* to do what morality requires of him: the freedom to subordinate the call of self-love to that of moral duty (i.e. to give absolute practical priority to the latter *vis-à-vis* the former) and to thereby instantiate what Kant calls a *good-will*, a will of the kind that alone can be called *morally good*. It is this freedom, this capacity to act in the world as a matter of moral duty, that constitutes our capacity for morality and makes us *moral creatures*, and it is hardly surprising that *being a moral creature* (someone's being capable of acting morally) be regarded as a morally significant feature of individuals, nor is it at all surprising that it be taken as the ground of *human dignity* in the substantive sense of the term. If you are a moral creature, i.e. a human being, then of course you are different in kind from non-moral creatures in general: you are raised above them along a strictly moral dimension and so are superior to them at least in that regard. Such is the basic meaning that we should attach to the word *humanity*, and such the basic meaning we should attach to the words *human dignity* substantively construed.

Again: humanity (i.e. a capacity for morality) makes human beings in general radically more important than non-human beings from the moral point of view. To be sure, and forgive me for the repetition, this is not the claim that human beings are radically superior to other animals *in ever morally significant way*: it is only to say that – as I am putting it – there is a special domain in the moral universe where only human beings belong, a

³⁴ By *responsiveness* here I mean a *capacity to respond*, not an *actual response*. Our 'humanity' is to be identified with our *capacity* to respond to the moral law, and our 'human dignity' (in the sense of our distinctive *moral standing*) has its basis in humanity *thus understood*. However, the *realisation* of our dignity as human beings is in fact based on our *actual response* to the moral law, and thus the idea that we have, not only a *moral standing* but also a certain *moral stature*. All human beings are by stipulation morally competent beings, and so *equals* in that sense, but some are (if they really are) *morally better* than others. I will say more about this in Section 4.2 (next section).

special position that human beings alone occupy, and an associated *role* or *function* that they alone are called to perform precisely because they alone have a *moral nature*.

The Kantian (substantive) conception of human dignity has now been characterised in its most essential features, but a problem arises already concerning the application of this concept. If it is open for human beings to make good or bad use of their moral capabilities, such that in principle (and speaking in simplified fashion) some of them may turn out *morally good* and some of them *morally evil*, then the question arises whether we are supposed to think of all of them as having the same measure of dignity, which sounds strange from the perspective of our everyday parlance (as we speak of human beings as acting in *dignified* and *undignified* ways, thus presumably entailing that some are *higher in dignity* than others).

4.2 Human Moral Equality and the Two Stages of Human Dignity:

It might seem problematic to say that ‘scoundrels’ and ‘saints’ are *equals in dignity*. We think of the saint as acting with a dignity (a stature) that is lacking in the conduct of the scoundrel, and this thought seems to contradict the claim (which I have been making in the name of ‘Kantian ethics’) that ‘human dignity’ and the associated idea of ‘human equality’ are to be understood in reference to our *capacity* for morality. To be clear, there is no complete and universal agreement on what the ‘true’ Kantian position is on these issues. For instance, some Kantians argue that human dignity is to be ascribed only to those human beings who we can regard as *morally good*, i.e. only to those who have “a commitment to actually obeying moral laws” (Dean, 2006, p. 246); others insist that no distinctions of dignity between different human beings (however moral or immoral they may be) should be made.³⁵ These

³⁵ The first position here is clearly a minority position, whereas the second one is perhaps only imaginary (although sometimes people may seem to be arguing for it). The reason I refer to these two ‘extreme’ views is only to introduce my own view (which I take to be in line with the ‘majority view’) as a compromise or middle ground between the two. Most (if not all) Kantians reject the second view because they agree with Kant’s claim that *morality itself* (the actual fulfilment of moral duties) has a certain dignity (G 4:439–440): they think (as they ought to) that morality is radically superior to immorality, that morally good people are (in the actual worth of their conduct) radically superior to the morally evil. Accordingly, they deny the claim that no

two positions are strictly speaking false, for they present us with a choice that we don't have to make; nevertheless, they contain the basic materials for a solution to the 'puzzle' we set ourselves to solve in this section. What we need to do here is actually very simple, namely differentiate the two kinds (or senses) of 'human dignity' that each of the views just presented emphasises and take them as two distinct *stages* of the one thing: 'human dignity'. Borrowing Oliver Sensen's terminology, I will name them: 'initial human dignity' and 'realised human dignity' (Sensen, 2011, p.153).

Here's the idea. There is a basic (or initial) kind of human dignity, or what could be called *initial human dignity*, pertaining to all human beings simply because they are 'human': this is of course their elevation (their uniqueness) as the only beings *capable* of morality (the only beings whose conduct is *subject to moral standards*) and their consequent equality as members of that domain. And then, of course, there is human dignity in its *realised* stage, or what could be called *realised human dignity*: the state of realisation of the human being as such (as a moral being) in the instantiation of a *morally good will*. Only those human beings who make a *good use* of their moral capabilities can be said to have dignity in the second sense, that is. The first (initial) notion of human dignity relates to the distinctive kind of *moral standing* that sets human beings apart from the rest of creation; whereas the second relates to the *moral stature* individual human beings give themselves on account of the moral quality of their actual conduct. Let me offer an analysis of this distinction and the central notions involved, which I think reveals important facts about the moral centrality of human dignity discourse as I am arguing should have it.

Let me pose our initial 'problem' again. Does the dignity of a scoundrel compare to that of a saint? Our answer should be yes. It follows from everything that has been said that in some important sense it *does* compare: the scoundrel and the saint are equals in moral standing, both having equal membership in the domain of *moral beings*. But if so:

distinctions of dignity *whatsoever* can be made between human beings. Those who reject the first view (the immense majority) do so because they think that all *human beings* are – quite irrespective of the moral quality of their conduct – worthy of what Stephen Darwall has called 'recognition respect' (Darwall 1977): (putting it simply) *respect for their humanity*. Accordingly, then, I think most Kantians would give their assent to the general view I am about to develop in the main text, deeming all human beings *equals in moral standing* but allowing for differences in their *moral stature*.

what happens when we point to the obvious *moral difference* between the saint and the scoundrel? Nothing happens. If you think about it, our judgements of moral stature putting human beings in the basic categories of ‘moral goodness’ and ‘moral evil’ are possible only under the assumption of their basic equality as beings *capable* of morality. This is very much like the need to first dignify claims or propositions as *candidates* for truth or falsehood – i.e. as being *truth-apt* – before judging them as either true or false. Surely we cannot assess the ‘truth-value’ of propositions (their being true or false) unless we first grant their *candidacy* (their aptness) for the ascriptions of the relevant values, and so raise them in standing above mere nonsense and other kinds of sentences that don’t purport to express truths. To put it in simplified fashion, truth-aptness (□) is presumed of both truths (+) and falsities (-). Similarly in our case, of course: surely we cannot assess the ‘moral worth’ of actions unless we first grant their *candidacy* for the relevant value ascriptions, which we do that by affirming the initial dignity of all human beings as such – as creatures whose conduct is subject to moral standards – and their consequent ‘superiority’ with respect to the non-human world. In simplified form: membership in the domain of ‘moral beings’ (□) is presumed of both the morally good (+) and morally evil (-). Just like both truths and falsities have to be regarded as equally *truth-apt propositions* (for otherwise they couldn’t be true or false), so both saints and scoundrels have to be regarded as equally *moral (morally competent) creatures*.

The purpose of this section is to gain clarity about our notions of *equality* and how they are used in relation to ascriptions of ‘human dignity’. Pushing the analogy above could help clarify this basic point even more. To begin with the epistemic case, the thought is that we are right to speak of truths and falsities as being *equals in epistemic standing*, since both are raised in kind (as truth-apt propositions) above, for instance, mere gibberish: nevertheless, this is not to deny the obvious fact that they are very much *unequal* when it comes to their *epistemic stature* (i.e. in their actual truth-value). There is a sense in which truths and falsities have ‘the same measure of dignity’, and another sense in which truths are ‘just superior’. It goes the same way in the case of human dignity, of course. We speak of the saint and the scoundrel as being *each other’s equals*, since they are both raised in *moral standing* above all entities lacking a morally rational nature (responsiveness to the moral law); nevertheless, it is obvious that they are very much *unequal* in their *moral stature* (i.e. in

the actual moral worth of their actions and/or character). There is a sense in which saints and scoundrels have ‘the same measure of dignity’, and another sense in which saints are ‘just superior’. In both cases, the initial notions of *equality* in the relevant standing work as *prerequisites* for the value ascriptions that follow, namely that some claims are true and others false, that some human beings are morally good and others morally evil.

My main goal in this **Part I** is to argue for ‘humanity’ (understood as a ‘capacity for morality’) as a dignity-conferring feature of individuals, a claim that comes with a commitment to ideas of *specifically human equality* (the moral equality of all human beings). What has been said in this section explains why the project of *human moral egalitarianism* is in no way threatened by the thought that some human beings are (if in fact they are) *morally superior* to others.³⁶

4.3 Morality as a ‘Dignity Term’

This has been quite a detour from our direct conversations with the Utilitarians. But we haven’t been wasting our time, I think, for now we have the building blocks to mount a convincing reply to the utilitarian critique of human dignity as we saw it above. My reflections in the preceding section are in some way internal to the Kantian story, but they reveal something important in relation to our more general argument. They show that *morality* itself is a ‘dignity term’, one of the terms characteristic of human dignity discourse, which of course means that if you want morality (and Utilitarians surely want morality) then you want human dignity (distinctively human dignity, I mean). Let me explain.

Our practice of moral assessment is axiological in nature, since there what we are doing is ascribing *moral valences* in the proper sense of the term (moral positivity and moral negativity) to the conduct and character of human beings. ‘Morality’ (moral goodness) is

³⁶ As I will explain in detail in **Part II**, the kind of respect in relation to human beings that figures as a fundamental moral requirement (the kind that is enjoined by Kant’s ‘formula of humanity’) is respect for human beings *as such*, i.e. simply as human beings (in their ‘initial human dignity’) and therefore quite irrespective of their moral stature. As I will also explain in **Part II**, there is a specific kind of ‘respect’ that is distinctively directed to *morally good* human beings, but this second kind of ‘respect’ is *not* understood to be a moral requirement but simply something that we do naturally or inevitably (if only as moral beings).

the most common name we use for moral positivity in this strict sense (+), and ‘immorality’ (moral badness or moral evil) the one we use for moral negativity proper (-).³⁷ These valences are clearly distinct from other philosophically interesting valences,³⁸ capturing as they do the *moral stature* of human beings (regarded in their capacity for ‘free choice’) and their actions (regarded ‘deeds of freedom’). Accordingly, the terms referring to these valences we may call ‘stature terms’. But if ‘stature terms’ are ‘dignity terms’ (which they obviously are), then *morality* itself (positive moral stature) is also by implication a dignity term: morality is the term we use in reference to the state of *realisation* of human beings as such (the realisation of their initial dignity as morally rational beings) and to the goodness of their conduct as in fact equivalent to that self-realisation. ‘Immorality’ or ‘moral negativity proper’ is nothing but the corresponding state of *failure*, of course.

More follows trivially. If (i) the main *explanandum* in moral philosophy is morality itself, and if (ii) morality itself is one of the terms of human dignity (a ‘dignity term’ in that sense), then (iii) everything we say in moral philosophy has a reference (at least implicit) to human dignity. The crucial premise here is the second one, for the first one is just plainly obvious. If you think about it, it is obvious (analytic) that *morality* itself is the main *explanandum* in moral philosophy. But the truth of the second premise in the argument should also seem obvious given what was said in previous paragraphs: morality itself (moral goodness) is obviously a matter of human dignity in the substantive sense of the term, i.e. if humanity (that which makes us ‘human’) is identified with our *capacity for morality* (our *capacity to be morally good*). Once ‘humanity’ has been defined in terms of the ‘capacity for morality’ we assume ourselves as having, it just becomes obvious that *morality* (that capacity for which constitutes our humanity) is a dignity term – one of the terms of human dignity discourse – referring to the realisation of the human being as such. And so we get the conclusion of the short argument at the top of this paragraph, which is the same as the

³⁷ Very often we talk about virtue (+) and vice (-) in this context, in reference to the same basic moral valences.

³⁸ Morality (moral goodness) and immorality (moral evil) are clearly distinct from, say, happiness and unhappiness, utility and disutility, prudence and imprudence.

opening claim of this **Part I**: *everything we say in moral philosophy assumes a substantive conception of human dignity of the kind that we find in Kant.*³⁹

Now, at this point we should really be wondering about the credibility of Utilitarianism as a view that supports (or is said to support) incredulity with respect to substantive notions of human dignity. For if the ‘Utilitarian critique’ is misguided (as we should by now be thinking that it is, at least as we saw it stated above), then the view that generated that critique in the first place (Utilitarianism) should be similarly misguided. In coming sections I will be offering an extensive argument against ‘Utilitarianism’ in its standard formulation (as most people understand the view in our present context), and as first preamble to that I would like to raise a few questions inviting reflection: Can we have a coherent moral philosophy without the idea that the *capacity for morality* confers a certain dignity (moral standing) on those who have it? If the answer is *no* (as I think it is), what follows about the nature of morality? And what about the credibility of Utilitarianism as a theory that – it is said – only has eyes for the capacity for happiness? I’ll address the basic concern that is expressed in these questions before going into the deeper argument about the nature and merits of Utilitarianism.

5. Interlude: On The Agent vs Patient Distinction

Advocates of human dignity have been accused of raising the moral standing of human beings on an arbitrary basis. But surely the Kantian view I have been developing is not guilty of *that*, since (as I have been trying to show) it can account very well for the dignity-conferring power of ‘humanity’ (someone’s ‘being human’ in the technical-philosophical sense) as a feature of particular individuals. Human beings are – by the stipulations the view – the only *moral beings* in the world, the only bearers of moral duties and the only

³⁹ The claim there, remember, was that moral philosophy is always to done under two related assumptions: (i) that human beings have *dignity* with respect to all non-human creatures, and (ii) that morality is all about the *realisation* of this dignity. Now we understand with perfect clarity what this claim means, and there is nothing mysterious or morally pernicious about it. I understand the suspicion that some people may initially have, of course: I understand why people become suspicious of human dignity discourse when it is most often used for things like opposing ‘abortion’ and justifying the extermination of animals for food.

candidate-bearers of that which constitutes *moral goodness* (as a distinct or *sui generis* kind of goodness).⁴⁰ The moral vocation is here identified with the distinctively human vocation, and the most basic assumption of all moral discourse is thereby supplied: the necessary reference to the agents (human beings) whose activity alone is subject to moral standards, in whose existence all moral realities are rooted.

But there is something very strange in all of this. If the assumption of human moral superiority (substantive or distinctively human dignity) is so basic that without it moral philosophy would lack a reference point and a subject matter, then we should find it very puzzling that people advancing moral claims of all kinds (as Utilitarians do) should stand in opposition to the making of this assumption. Let me suggest a solution to this apparent puzzle. If ideas of human moral superiority can be *misused* (i.e. used for evil purposes), and if – as I think is the case – they have in fact been misused, then it might be *this* phenomenon that Utilitarians (and others who we may identify as ‘critics of human dignity’) have been reacting so strongly against. I mean: maybe it is not ideas of substantive human dignity *themselves* that are problematic in Utilitarian philosophy, but only appeals to those ideas, as premises, to argue for the objectification (or less than full moral appreciation) of animal life as such, which is what Kant and other champions of human dignity seem to have been doing.⁴¹ Kant, for instance, is naturally read as saying that by virtue of human moral superiority we are entitled (and even required) to regard non-human animals as ‘things’ and treat them as ‘mere means’ (see **Part II**, Section 1.); and what I am suggesting is that the Utilitarian critique of anthropocentric ideas (in relation Kant and people arguing along

⁴⁰ Again, the stipulation here is not that all the members of the ‘human species’ (the biological category) should be counted as moral beings. Rather, it is simply that those beings who in fact count as moral beings (because they have a capacity for morality) we should count as ‘human beings’ in the technical-philosophical sense of the term that is relevant in discussions of the kind that we’ve been having.

⁴¹ I wonder how Utilitarians like Singer would have reacted to substantive ideas of human dignity if those who made them central to their philosophies (Kant in particular) had made it clear that (i) non-human animals are not meant to be reduced to ‘things’ but must instead be regarded as ‘persons’ (even if not as ‘moral subjects’), and (ii) that non-human animals are, no less than we are, proper objects of the kind of benevolence (kindness) that morality fundamentally requires. Feel free to disagree, but my hunch is that there would be no such thing as a Utilitarian (or in any way ‘sentio-centric’) critique of human dignity.

similar lines) should be interpreted as agreeing with the premise (human moral superiority *as such*) but disagreeing with the conclusion being derived from it.⁴²

If we understand the Utilitarian critique in this way, i.e. as not objecting to ideas of human moral superiority themselves but only to *misuses* of them in the objectification (or less than full moral appreciation) of animal or sentient life, the initial puzzlement seems to vanish completely.⁴³ Again, what seems puzzling in the ‘Utilitarian critique’ as we say it above (in its two initial premises and their respective conclusions) is the apparent denial of a moral standpoint characterising the human – as opposed to the animal – condition. But that is really *not* what utilitarians mean to be doing, I don’t think. If some of them see themselves as doing that, I invite them to reconsider: our arguments in defence of our ideas of *animal dignity* and *animal equality* (as pertaining to the domain of intended beneficiaries of morality) shouldn’t depend on denying human uniqueness, since the whole argument is precisely about *that*: the call (the duty, the mission, etc.) to love all sentient beings without distinction (under a principle of equality) is something *distinctively human*. In fact, you don’t have to agree with my proposed ideas of ‘animal dignity’ and ‘animal equality’ to appreciate that in putting them forward for consideration I am trying to say something *morally substantial*, something about the human (moral) condition, namely that in the eyes of human beings (in our moral eyes) *all animals* figure in a certain way. Anticipating a little, the point here is that there is a *moral duty* of kindness or benevolence that puts us (moral beings) in relation to sentient beings *in general* as intended beneficiaries of our own activity, a relation in which the little kid figures in the same as the old lady, the first minister the same as the nomad, and the human being the same as the pig, and the cow, and the chicken.⁴⁴ This kind of *animal egalitarianism* is a *moral idea*, an idea about how individuals (persons) are to be regarded from the moral (distinctively human) point of view.

⁴² This means that they object to the *principle* (the so far unstated premise in Kant’s argument) by which the conclusion is derived from the premise. My argument in relation to Kant’s ‘animal ethics’ (which does provide a connecting principle that can be challenged) constitutes a large portion of **Part II** below.

⁴³ It is not puzzling, I think, that people would (i) oppose attempts at objectifying the ‘animal condition’ and (ii) try to defend the dignity (special moral standing) of all sentient beings. That is as it should be.

⁴⁴ It is part of my argument in **Part II** that all animals count as persons for moral purposes. Given this thesis and the other developments made in **Part II**, the maxim just stated in the main text should be read as follows:

Again, whatever your take is on ‘Utilitarian’ (or in any case ‘sentio-centric’) ideas, you can appreciate that in putting them forward for consideration I need to presuppose a certain conception of human dignity (in the substantive sense of the term) giving me a moral reference point, making sense of the very place where I stand on, and the audience I am addressing, in thinking and talking about these issues. In particular, the duty of kindness towards all sentient beings that I think *we* have is meant to be a *moral duty*, and as such it presupposes *our* condition as the relevant duty-bearers (the bearers of this particular moral duty). Thus, if Utilitarians believe (as of course they do) in the existence of a *moral duty* of universal kindness or benevolence,⁴⁵ and if they think (which of course they also do) that the burden of this duty falls *exclusively* on human shoulders, then surely they have to acknowledge that *in our capacity for morality* we (human beings) are radically superior to non-human animals. There seems to be no way around the fact that Utilitarians – however ‘radical’ they may be in their Utilitarianism – *have to* presuppose ideas of human *moral* superiority if only to place the moral burden on the right shoulders. Utilitarians are no different than anybody else in this regard, as for instance Onora O’Neill has acknowledged in her description of a *necessary logo-centrism* (O’Neill, 1998). The success of the ‘Utilitarian’ cause – the cause of defending the notions of ‘animal dignity’ and ‘animal equality’ – cannot be predicated on denying that which is distinctively human: the Utilitarian cause is a distinctly human cause, a project for morally rational beings.

Let me make the same point from a slightly different angle. The argument so far has been mainly that Utilitarianism is best read as being *implicitly* committed to ideas of substantive human dignity. However, there seems to be some *explicit* (if only perhaps inconsistent) support as well. If you think about it, the affirmation of human moral superiority is normally taken as a *premise* in the project of ‘utilitarian animal ethics’ as we have it already in the literature (even in the work of Singer himself, who therefore is wrongly read as denying any special place for human beings in the moral universe), the central case of which goes as follows: non-human animals are *not* moral agents, granted, but this doesn’t

all persons (insofar as they are capable of happiness, and quite independently of their ‘moral credentials’) are equal in their condition as *intended beneficiaries* of the moral enterprise.

⁴⁵ I assume the truth of this claim in the thought that the duty in question is the quintessential Utilitarian duty.

condemn them to having the status of a ‘mere means to human ends’, since they can still count as *moral patients* with respect to us, meaning that we (moral agents) can still have *duties to them*.⁴⁶ The truth seems to be that Utilitarian philosophy is *already* committed to ideas of *substantive human dignity* (human moral superiority) and *specifically human equality*, the only qualification being that these ideas have to be understood as pertaining to the ‘first side’ in the moral equation as it were: the placing of *moral burdens* on the relevant shoulders. If understood as belonging *also* to the ‘second side’ of the moral equation, i.e. to the moral patient side (which for Utilitarians tends to take the form of an *intended beneficiary*), anthropocentric ideas are completely out of place. Again, it is the idea that ‘humanity’ fixes the domain of ‘moral patients’ that Utilitarians mean to be taking exception to, not the mere affirmation of the kind of human *moral* superiority that – as I have explained – is necessarily assumed in all moral theorising. What Utilitarians are telling us is that when it comes to the issue of moral patients (*intended beneficiaries* in particular), it is the dignity of *all animals* – human and non-human alike – that has to be honoured, which we do (philosophically speaking) by saying that morality is something that we do *for the benefit (the happiness) of all sentient beings*.

Kantians can respond to the initial ‘Utilitarian critique’ as we saw them doing already, pointing to ‘humanity’ (as a ‘capacity for morality’) and its paramount moral significance: its being the ground of all moral realities, for instance. But if – as I am suggesting – Utilitarians can come to agree on this point, situating ‘humanity’ in the picture as *the* moral reference point; and now that the target of their critique has been qualified

⁴⁶ Lori Gruen, for instance, has put the point as follows: “The moral universe, as it were, thus can be said to have two levels: a level that contains moral agents, actors who are responsible for doing the right thing; and a level that contains moral patients, those to whom right or wrong actions are directed, but who may or may not be moral agents themselves” (Gruen, 2011, p.61). Paola Cavalieri has explained the distinction between moral agents and moral patients by saying that: “If the moral agent is a being whose *behavior* may be subject to moral evaluation, the moral patient is a being whose *treatment* may be subject to moral evaluation” (Cavalieri, 2001, p.29). She also thinks that “the criterion of consciousness remains the only plausible candidate with respect to the access to the status of moral patient” (Cavalieri, 2001, p.40), and on this basis includes all animals (all sentient beings) in that domain. The domain of ‘moral agents’ remains exactly as Kantians would have it, taking it for granted that “only rational and autonomous beings can be morally responsible” (Cavalieri, 2001, p.28).

accordingly, it seems as if Kantians who reduce animal or sentient life to nothing more than a 'mere means' (something with 'price' instead of 'dignity') are either (i) putting themselves in a difficult position or (ii) called to say much more than they have done so far, to explain to us exactly why we should believe as they do. Why is it that *only* human beings count as moral patients? More importantly still, why should humanity (the capacity for morality) be taken to fix – or be in any way material to fixing – the domain of direct moral considerability (and in particular the domain of intended beneficiaries of morality)? Addressing questions of this kind demands introducing a lot of nuance in the argument, and actually they are material for **Part II** below (where my argument for the dignity of sentient beings in general consists mostly of showing how it fits into a basic 'Kantian' – including 'Contractualist' – framework positing various irreducible intra-personal and inter-personal moral relations, some of which are such that only human beings count as direct participants).

The basic goal of this section has been to qualify the Utilitarian (sentio-centric) critique above as objecting, not to ideas of human moral superiority *as such* but only to the *misuse* of them in the objectification or less than full appreciation of animal or sentient life: the objectification of individuals, human and not, *in their capacity for happiness*.

Now, all of the above might be nice to hear, but it seems to me that Kantians have grounds to still remain sceptical of the story I have been telling on behalf of the Utilitarian view, especially as one deserving the 'Utilitarian' label. If Utilitarians are telling us that they agree on 'humanity' being essential as a moral reference point, i.e. as providing the first (properly moral) side in the moral equation, then surely this means that they also have to agree with making 'humanity' (human life) the very ground of all moral realities (precisely because it fixes the properly *moral* side in the moral equation); and yet, if you read the accounts of the nature of *morality* that Utilitarians typically offer, and if you pay attention to the conceptions of moral agency they typically work with, most (if not all) the work seems to be done by felicitic (i.e. happiness related) considerations. For instance, they typically tell us that morality is a matter of *utility*, of actions having the best *consequences over the well-being of sentient creatures* (be that in actuality or in expectation). It's as if Utilitarians were giving 'humanity' its proper due with one hand, only to immediately dispossess it with the other. And so I ask: Why should we bother talking about distinctively human dignity if the actions

of cats, and trees and rocks also have consequences of the kind that are said to constitute the *morality* (the moral value) of human activity? Why should we bother if moral valences in the strict sense of the term (*morality* and *immorality*) depend on something that doesn't belong in human life *as such*?⁴⁷

It is not a secret that Kantians take issue with Utilitarian conceptions of 'morality' as we know them and the related conceptions of 'moral agency' they seem to be working with. The way Kantians typically regard the matter means that we should be questioning whether Utilitarians (insofar as we see them trying to reduce 'morality' to 'utility' or 'conduciveness to everyone's happiness' in some form) are even engaging in *moral* talk. Maybe it would be a little hasty for a Kantian to insist on this suggestion, since sometimes it happens that *false reductionisms* (and I agree that the one just mentioned – the reduction of morality to utility – is a textbook example of that kind of mistake) have some important truth lying behind them which people are failing to spell out correctly and which may escape people who don't even try to go there.⁴⁸ I submit that Kantians should hold their judgement over the possibility of a credible Utilitarian story about the nature of morality at this point: I would invite them to first hear what I have to say about what Utilitarianism

⁴⁷ This may sound as already misrepresenting Utilitarianism as an ethical theory, for it is by no means clear that the view necessarily involves a commitment to morality depending on anything other than the intrinsic qualities of human volition. Consider this line by Mill (1861/2003, p.196): "The morality of the action depends entirely upon the intention – that is, upon what the agent wills to do. But the motive, that is, the feeling which makes him will so to do, when it makes no difference in the act, makes none in the morality: though it makes a great difference in our moral estimation of the agent." I don't want to read too much into this statement by Mill just yet, but it should at least generate doubts on Utilitarianism actually reducing the moral qualities of human activity to its felicific properties considered in a certain light (without of course having to deny that there are all kinds of statements by Utilitarians suggesting otherwise, as we will see shortly in the main text). What I mean to do here is warn you that the argument I am about to present against 'Standard Utilitarianism' ('Consequentialism' + 'Welfarism') is not an argument against what I regard as *Utilitarianism* in its correct formulation. **Part II** is my attempt to develop and defend what I regard as the real thing (thus hopefully redeeming myself for the occasional sweepingness of my claims in the argument shortly to come).

⁴⁸ That is exactly what I think happens in this case. For my diagnosis of false reductionisms in which both Utilitarians and Kantians tend typically to fall into, see Sections 3. to 5. of **Part II** (culminating in long footnote in subsection 5.5, which condenses most of the general argument).

amounts to (which I do mostly in **Part II**), if only for us to be clear on what exactly it is that they want to reject.⁴⁹ What Kantians have to do at this point, I think, is *criticise* those ‘Utilitarian’ conceptions of morality and moral agency that they reject as false on the basis of their own positive conceptions; more importantly still, they are called to point to those aspects of their own conceptions which they take to be crucial, so that the fruitful conversation that I am proposing can move forward.

For the moment, I need to deal with the undeniable fact that ‘standard’ Utilitarian conceptions of morality and moral agency appear to be in tension with their Kantian counterparts. The general argument for ‘Kantian ethics’ (i.e. for what I take to be the essence of Kantian ethics) will emerge out my critique of Utilitarian conceptions of the ethical as we usually find them formulated in the literature.

6. Against ‘Standard Utilitarianism’ (For Kantian Ethics)

I have been developing the Kantian doctrine of human dignity *vis-à-vis* alleged (and progressively vanishing) Utilitarian incredulity, touching mainly on conceptual and definitional matters with the aim clarifying and qualifying my claims. However, I have barely touched on more substantive issues concerning the nature of morality itself, upon which everything depends.

The Kantian makes human dignity revolve around the ‘capacity for morality’ that separates human beings in kind above non-human entities in general; but of course anyone can accept this thought and propose his own account of the nature of *morality*, which may not be Kantian in the least. Utilitarians, for instance, could well agree to the elevation of humans beings as moral subjects, and yet make morality depend on the felific properties of human activity: its consequences and consequential tendencies with respect to the well-being of sentient creatures in general. Such view could hardly be called *Kantian*, and by the same token it would also seem strange to put this label on the corresponding account of human dignity. If humanity (the feature of someone’s *being human* is the relevant sense of

⁴⁹ In **Part II**, I will be arguing that Utilitarianism should not be styled as a version of ‘Consequentialism’, but instead as a kind of *Ethical Agapism*.

the term) didn't constitute an *independent* source of value; if the human condition didn't come with a capacity to realise something of *intrinsic value*, i.e. something whose value could not possibly be reduced to the value of something else belonging to a different condition (e.g. happiness, which belongs to the 'animal condition'), then it is hard to see how human beings could *as such* have dignity, or (in other words) how there could be such a thing as *substantive* human dignity: the 'value' of human beings (if I may put it this way) would be reduced to the one that attaches to them as bearers and/or causes of happiness. Besides, the issue is not whether such a view would deserve to be called 'Kantian', but instead whether it is *true*, and I submit that it is not true.

I am of course referring to the 'standard' utilitarian view (the one that is dominantly taught in universities), which explicitly reduces all 'goodness' (even presumably *moral goodness*) to 'felicific goodness' and as such can be seen as destroying our intuitive categories: it destroys morality (moral value) *properly defined* (i.e. as residing in the internal activity of a moral being, and as consisting in activity that is done *from duty* or in any case *from a distinctively moral motive or maxim*) and, also by implication, human dignity *substantively understood* (i.e. understood as belonging to human beings on account of their *humanity*). Kantians are in a very different position in relation to these issues, as they seem to have what it takes to provide a non-reductive conception of morality and so also a credible account of the kind of human dignity we are looking for: the kind of dignity that belongs to human beings *in exclusivity*, the kind of dignity that cannot be reduced to the purported 'animal dignity of the human being'.

From this point onwards, what I will do is develop my case for a recognisably Kantian conception of *morality*. However, the argument is meant to emerge from a 'Kantian critique' of Utilitarianism as we know it in dominant narrative in philosophy (i.e. as we have it in the books and teach it in universities). The 'standard utilitarian formula' – as I am going to call it – has portrayed *Utilitarianism* as the conjunction of 'Consequentialism' and 'Welfarism': as the view according to which the moral qualities of actions are a function of the *relevant consequences*, coupled with the view that the relevant consequences in that equation are the ones concerning the *well-being* (or happiness) of sentient creatures in general. Both those constitutive views (both 'Consequentialism' and 'Welfarism') make demonstrably false claims, I argue, and thus Utilitarianism is also by implication

demonstrably false.⁵⁰ Here, again, I am referring to Utilitarianism in its ‘standard formulation’, leaving it open whether there is some better way of construing the view, which I argue there is. The following two sections tackle the ‘Consequentialist’ side of the equation, and the one after provides an extended argument against ‘Welfarism’.

6.1 Against ‘Consequentialism’ – Defining the View

‘Consequentialism’ has been given definitions that make it a view in *moral metaphysics* or *moral ontology*, i.e. a view that purports to account for the *nature of morality*, and it is in that capacity (as a view in moral metaphysics) that ‘Consequentialism’ will be a target for me in this section.⁵¹ In his entry to the Stanford Encyclopedia of Philosophy, Walter Sinnott-Armstrong offers one such definition, which I take will take up as my *initial definition* in preference to other definitions that have been proposed (Sinnott-Armstrong, 2019):

⁵⁰ To anticipate the crux of my argument against ‘Welfarism’ (the thesis that happiness or well-being is ‘the only intrinsic good’), this view is false not because there are kinds of outcomes other than felicitic outcomes which we should morally pursue in the manner that (as Utilitarians think of it) we ought to pursue felicitic outcomes. The problem with ‘Welfarism’ is rather that it is blind to the ‘intrinsic moral value’ that attaches to the good-will *as such* (regarded *as cause*, not as consequence).

⁵¹ Some people understand ‘Consequentialism’ as the view according to which morality is all about the ‘promotion of the good’, issuing no ‘side-constraints’ (no constraints or prohibitions above and beyond the ones that the ‘the promotion of the good’ involves) and no ‘special prerogatives’ for the agent to give disproportionate weight to his good *vis-à-vis* the good of others (Scheffler, 1982; Kagan, 1992; Cummiskey, 1996). I won’t say much (if anything) about ‘Consequentialism’ under this definition, partly because I need definitions with more of a focus on how ‘Consequentialism’ purportedly depicts the *nature* (in a descriptive sense) of moral realities and the moral qualities of human activity, and also partly because (depending on how the view be developed further) I may just be a ‘Consequentialist’ under this particular construal (as I think there are no moral prohibitions above and beyond the ones that are grounded in *agape*, and so no prohibitions that don’t exist for everyone’s happiness, nor do I recognise the existence of ‘special prerogatives’). Perhaps my major qualm with respect to dominant theorising in this area is that all specific moral prohibitions being grounded in *agape*, and so being constitutive of the ‘promotion of the good’, as I think they are, in no way entails that these prohibitions are anything less than *absolute prohibitions*.

Consequentialism, as its name suggests, is simply the view that *normative properties depend only on consequences*. [...] This general approach can be applied at different levels to different normative properties of different kinds of things, but the most prominent example is probably consequentialism about the *moral rightness of acts*, which holds that whether an act is morally right depends only on the consequences of that act or of something related to that act, such as the motive behind the act or a general rule requiring acts of the same kind.⁵²

The first thing to say about this definition is that it does a good job at capturing the theoretical ambitions of the view: as defined, ‘Consequentialism’ is a daring attempt to account for the ‘normative properties’ of actions (and of other things perhaps) by reference to the one thing that – as they see it – *ultimately* matters about them, namely their *relevant consequences*. For Utilitarians the ‘relevant consequences’ are those that pertain to the happiness or well-being of sentient creatures (because they see themselves as *welfarists* about value), but Consequentialists endorsing other axiologies will identify different consequences as the relevant ones to consider.⁵³ I will keep referring to ‘Consequentialism’ having the generic view in mind, but please forgive me if at any point I lapse into addressing Utilitarians in particular.

So far so good, but we cannot move forward without first touching on at least two other points concerning our initial definition. Please note that this definition speaks of ‘normative properties’ as that which is being explained by the view, which is potentially misleading given that the normative domain is wider than the moral domain, including as it does things like aesthetics and epistemology. We can rectify this problem while still following the path of our initial definition by noting that ‘Consequentialism’ is a view

⁵² Both emphases are mine.

⁵³ This is a familiar point in the literature. The only thing to add is that instead of talking about ‘relevant consequences’ we could speak of the ‘value of consequences’. You can think of actions as having many kinds of consequences, some of which are ‘of value in themselves’ and some of which are not. The ‘value of consequences’ reduces to the value of those consequences in particular that are of value in and of themselves, and we can speak of the latter as the ‘relevant consequences’.

purporting to account for *moral qualities*, which it does by positing a metaphysical dependence of all such qualities on the relevant consequences (whatever those are).⁵⁴

One last thing to note is that the definition above contains an idea that has come to be known as *Global Consequentialism*: the idea that ‘consequentialist evaluation’ can be applied “at different levels to different normative properties of different kinds of things” (Sinnott-Armstrong 2019). We can put this idea to the side for the moment. After all, Consequentialists typically restrict the domain of application of their principles to the few things that figure in our common moral discourse as fitting objects of *moral evaluation*: the actions and underlying intentions, motives and rules of human beings.⁵⁵ Our initial definition fairly captures this point when it says that the “most prominent example is consequentialism about the moral rightness of acts” (Sinnott-Armstrong 2019), although the truth is that consequentialists also often speak of the ‘value of actions’ (e.g. Scarre 1996 p.101), the ‘moral value of actions’ (e.g. Crisp 1997, p.140), the ‘moral status of actions’ (e.g. Driver 2012, p.86), and the ‘morality of actions’ (e.g. Mulgan 2006, p.55).

The kinds of views that Consequentialists typically propose, then, are ones that focus primarily on the evaluation of *actions* (the actions of the relevant agents), and different terms are used to pick out the ‘moral qualities’ that actions are identified as having: they call actions ‘right’ and ‘wrong’, ‘good’ and ‘bad’, all of this – they say – on account of (considerations of) their relevant consequences. We can then take all these points into account and modify our definition of ‘Consequentialism’ accordingly, thereby crafting a *second definition* of the view that we can take up as our own: ‘Consequentialism’ – we can now say – is the view according to which the *moral qualities of actions* (i.e. of the relevant actions, i.e. human actions) ultimately depend on (considerations of) their relevant consequences.

We are not quite there yet. It seems to me that our second definition is a decent enough version of the ‘standard definition’, but it is still vague in one important respect.

⁵⁴ This is not a point against Sinnott-Armstrong’s treatment of this issue, as he explicitly qualifies his definition of Consequentialism as a view about the *moral* properties of actions (Sinnott-Armstrong 2019).

⁵⁵ They don’t necessarily deny that we can appreciate the ‘utility’ of all things that have a ‘utility’. What they deny – I assume – is that evaluations targeting things other than the distinctive activity of rational agents count as *moral* evaluations.

We are about to see arguments against ‘Consequentialism’ *as a theory in a certain capacity*, and we absolutely need to conceive of it as a theory in *that* capacity. My arguments against ‘Consequentialism’ take it as a theory that purports to account for the *nature of morality*, and even though talking about the ‘moral qualities of actions’ seems to do justice to that, the case is that it may not. Well, this depends on what ‘moral qualities’ we have in mind and how we define those qualities. If the qualities of actions that we have in mind are things like their ‘rightness’ and ‘wrongness’, or even (if you wish) their ‘moral rightness’ and ‘moral wrongness’, then (depending what we mean by these properties) we might be losing sight of the real thing – we might be losing sight of the *morality* of actions.

The concern just raised is based on the Kantian distinction – that I endorse – between the ‘legality’ and the ‘morality’ of actions, a distinction Kant illustrates in his famous portrayal of a shopkeeper in the *Groundwork* (G 4:397). As you probably know, Kant’s shopkeeper is a person who does *the right thing* (charging the same price to all his costumers) but does so on *non-moral grounds* and so fails to give *moral worth* to her actions. In terms of the original distinction, the person here acts *legally*, but nevertheless fails to act *morally*. We can also think of the person who adopts a vegan or plant-based diet (excluding all animals from her menu) but does so on *non-moral grounds*: in some sense this person does *the right thing* and so acts *legally* (with regards to her abstinence at least), but *morality* is nowhere to be found there. Now, if this distinction is real and ‘morality’ is not to be found in simply *doing the right thing* (in the mere *external conformity* of actions with moral commands, which only gives us ‘legality’ or ‘rightness’), then we should indeed find it problematic to take ‘Consequentialism’ as a theory about *that*. I mean, if we really want to see ‘Consequentialism’ as trying to account for the nature of *morality*.

The picture has become messy before us even getting going with the argument. To make things simple, let me propose one final amendment to our second definition above. Since we need ‘Consequentialism’ to be a theory about the nature of *morality* (the real thing as opposed to mere legality of conduct), I see no option but to *make it so* and see what happens. Our *final definition* then has it that ‘Consequentialism’ is the view according to which the *morality of actions* depends entirely on their relevant consequences. Now, if it turns out that ‘Consequentialism’ aims to account for the nature of morality (the real thing)

and fails in that attempt, then all the worse for it. By construing ‘Consequentialism’ in this way we at least give it the dignity of counting as a false moral philosophy.⁵⁶

6.2 Against Consequentialism – The Argument

Let me pose the question. What is wrong with ‘Consequentialism’ as a theory aiming to provide us with our basic moral framework? Simple: it tells us that the morality (moral value) of actions is a function of their relevant consequences, when the truth is the complete opposite of that: actions are ‘moral’ or ‘immoral’ only on account of their relevant *causes* – only on account of the motives or the maxims from which they are done. Morality is irreducibly a matter of *attitude* on the part of the moral agent herself, not a matter of universal ‘utility’ or ‘expediency’: it is a matter of *acting on good motives or maxims*, not a matter of producing (or even promising to produce in the ‘expected value’ sense as applied to human conduct) the best consequences. We have come to associate these claims with Immanuel Kant, the ‘anti-consequentialist’ *par excellence*, although the truth is that many (if not most) philosophers in history have defended some version of this view (many of them portraying it as the view that all human beings of themselves already hold).⁵⁷ Now,

⁵⁶ I acknowledge that nobody regards ‘Consequentialism’ as a complete moral theory. What my argument is meant to show, then, is that any complete ‘Consequentialist’ theory (any theory making the morality of actions depend in any way on the value of its consequences, whatever those are) is false.

⁵⁷ Kant portrays this and closely related thoughts as part and parcel of our ‘common moral cognition’. But it’s not only Kant who thinks this. We find basically the same view in philosophers historically more closely associated with the ‘Utilitarian tradition’, e.g. Adam Smith, whom I’d like to quote on the particular (1759/1976, p.93): “The only consequences for which he can be answerable, or by which he can deserve either approbation or disapprobation of any kind, are those which were someway or other intended, or those which, at least, show some agreeable or disagreeable quality in the intention of the heart, from which he acted. To the intention or affection of the heart, therefore, to the propriety or impropriety, to the beneficence or hurtfulness of the design, all praise or blame, all approbation or disapprobation, of any kind, which can justly be bestowed upon any action, must ultimately belong. When this maxim is thus proposed, in abstract and general terms, there is nobody who does not agree to it. Its self-evident justice is acknowledged by all the world, and there is not a dissenting voice among all mankind. Every body allows, that how different soever the accidental, the unintended and unforeseen consequences of different actions, yet, if the intentions or affections from which they arose were, on the one hand, equally proper and equally beneficent, or, on the

given that I find Kant particularly helpful in discussions of this kind, I will take the liberty to speak of the ‘Kantian thesis’. To repeat: *morality* and *immorality* (moral valences proper) depend entirely on the *relevant causes* of actions.

If the ‘Kantian thesis’ is true, ‘Consequentialism’ is false (at least as an ethical theory). That much is clear. Now, one thing to note about the ‘Kantian thesis’ is that it follows from a certain *axiological view* that I also take the liberty to designate with the name ‘Kantian axiology’ (which, to be clear, is not to say that this *kind* of axiology is only present in Kant). The problem with ‘Consequentialism’ (as a general framework upon which to build a moral theory) is ultimately to be explained with reference to this axiological view, which might at first sound paradoxical given that ‘Consequentialism’ is normally understood as being neutral between different axiologies. Now, given my claims here, I should be able to make this apparent paradox disappear, which I do as follows: if ‘Kantian axiology’ undermines ‘Consequentialism’ because it generates a thesis about the nature of *morality* (as of itself an axiological item) that makes the consequences of actions immaterial to their morality (the Kantian thesis), then the truth is that ‘Consequentialism’ is not so axiologically neutral after all.⁵⁸ But how does ‘Kantian axiology’ generate the ‘Kantian thesis’? And what does ‘Kantian axiology’ say in the first place?

One of the essentials of ‘Kantian axiology’ is the claim that the *good-will* (the morally determined will) is *intrinsically good*: the good-will is good *in itself* and as such has no need of something else (e.g. happiness) conferring goodness upon it for it to be good. Kant thus

other, equally improper and equally malevolent, the merit or demerit of the actions is still the same, and the agent is equally the suitable object either of gratitude or of resentment.”

⁵⁸ The reason why Consequentialism has come to be understood as neutral between different axiologies is that reflection on the *distinctive* valences attaching to the *relevant causes* of actions (relevant from the perspective of moral judgement properly understood) has been avoided. They attach valences to causes (even to the *relevant causes*, i.e. to those items that figure as ‘moral evaluands’ in our common moral discourse) on account of their ‘relevant consequences’, to be sure (which they are right to do insofar they mean to be accounting – in the case of ‘Welfarists’ at least – for the ‘beneficial character’ of human activity); however, insisting on these ‘consequentialist valences’ as the only ones in the picture – which they seem to do because otherwise they wouldn’t call themselves ‘Consequentialists’ – is to completely omit those *distinctive* valences I am talking about (*moral valences* in the proper sense of the term): morality (+) and immorality (–), which are always predicated of the causality of the human will *as such*.

famously compares the good-will with a jewel shining its own light (G 4:394).⁵⁹ Now, it is not just any kind of goodness that attaches to the good-will in this intrinsic kind of way: only the kind of goodness we call ‘moral’ does. The good-will is thus *intrinsically morally good* – it has (we might say) *intrinsic moral goodness*. Now, it is exactly this piece of Kantian axiology that generates the ‘Kantian thesis’, and my main task in this section is to explain how this transition happens. I will also try to defend the axiological claim itself, but that will be in the next section (where I offer an extended argument against ‘Welfarism’ on the basis of an argument for ‘Kantian axiology’, thus also completing my argument against ‘Consequentialism’).

So, if you take the Kantian affirmation of the *intrinsic moral value* of the good-will, and put it together with the claim that the good-will counts as a *cause* of actions (in the relevance sense of *counting as a cause* in this context),⁶⁰ then we start to get very close to the ‘Kantian thesis’. Let me make this into an argument. (1) If ‘moral goodness’ attaches *intrinsically* to the good-will (the morally-determined will), and (2) the good-will counts as a *cause* of actions (in the relevant sense of ‘counting as a cause’ in this context), then the conclusion follows that (3) ‘moral goodness’ attaches *intrinsically* to something (the good-will) that counts as a *cause* of actions in the relevant sense. Premise (1) will be defended in the next section, but we can take it for granted for the sake of the argument here. Premise (2) is more important for us at this point, since it gives us the thought of the good-will (that which gives us moral goodness) as something that counts as a *cause* of actions in the relevant sense. And this is important because: if one has the further thought – the truth of which

⁵⁹ I will amply discuss the relevant passage (Kant’s ‘jewel passage’ in the *Groundwork*) in the next section.

⁶⁰ I’m speaking of the good-will as ‘counting as a cause *in the relevant sense*’ rather than just as ‘counting as a cause’ (*simpliciter* as it were) mainly because I don’t mean to deny the metaphysical status of the good-will as a *consequence*, as something that is brought about by certain causes. I don’t find it crazy to think that the good-wills in the world (if there are any) count – at least in some sense – as *consequences* with respect to the ‘creation of the world’, for instance, and even perhaps with respect to the *choices* on the part of the relevant subjects to act morally (although those choices would themselves have to count as constitutive of good-willing). What I mean to indicate is simply that it is an *assumption* of the practice of moral assessment that we must be targeting the relevant kind of *causality* of the relevant kind of *agents*: we must be targeting human conduct (the conduct of morally rational beings) regarded as *free*, i.e. as *responsive* to moral considerations and therefore ultimately to the moral law.

seems obvious – that ‘morality’ is all about the realisation of ‘moral goodness’, i.e. if ‘morality’ and ‘moral goodness’ are two names for one and the same positive valence (moral positivity proper),⁶¹ then one has made it all the way to the ‘Kantian thesis’: morality (‘moral goodness’ obtaining) is something that depends on the relevant cause of actions (the good-will) being instantiated by the relevant agents. Morality (the real thing) depends entirely on the *cause*, not a single bit on the consequences.

This argument against ‘Consequentialism’ as a theory in moral metaphysics (or moral ontology) is meant to be *conclusive*, but I am yet to defend its crucial premise: the piece of ‘Kantian axiology’ that generates the ‘Kantian thesis’. The claim, that is, that the good-will (‘moral virtue’ if you will) is *intrinsically good* in the relevant sense (the only thing that has ‘intrinsic moral value’ and so the only thing capable of giving us *morality of actions*). Everything hinges on whether this piece of axiology holds true, and I will be arguing for it under the rubric of ‘Kantian axiology’.

6.3 Against ‘Welfarism’ (For ‘Kantian Axiology’)

The Welfarist theory of value that Utilitarians typically endorse has two claims: (i) that happiness (well-being)⁶² has ‘intrinsic value’, and (ii) that happiness is *the only* thing of ‘intrinsic value’. The first claim is true, the second false. I reject (ii) on the basis of the ‘Kantian thesis’ as characteristic of what I am calling ‘Kantian axiology’: the claim that the

⁶¹ Similarly with ‘moral evil’ and ‘immorality’, of course.

⁶² In line with conventional Kantian and Utilitarian practice (without denying the exceptions), I’m using ‘happiness’ and ‘well-being’ as synonyms. I take them to stand for the specific kind of *benefit* that is the object of love (love as ‘benevolence’ or ‘kindness’, that is), which we can name the *substantively beneficial* in order to distinguish it from other senses of the beneficial (e.g. the beneficial in the *functional* sense of the term: that which constitutes or contributes to the well-functioning of an entity that by nature has that function). In any case, it should always be clear that we are talking about goodness (benefit) in the *well-being sense*, i.e. about the *felicific* (happiness), which obviously doesn’t apply to all things that have functions. A car, for instance, no matter how well it performs its functions, cannot be called *happy* (at least not literally).

‘good-will’ (the ‘morally-determined will’) is *intrinsically good*. I will argue for this claim – thus undermining (ii) – after giving a basic explanation for why (i) holds true.⁶³

6.3.1 The Intrinsic Value of Happiness

Let me then start by arguing for happiness being something of ‘intrinsic value’ (a claim which I consider essential to what I call ‘Kantian axiology’, as it will soon become more explicit).⁶⁴ The claim here is that happiness is something that is *good in itself*, something that is good simply in virtue of what it is, something that has no need of an external thing conferring goodness upon it for it to be a good thing, something that contains the source of its own (distinctive) value, a.k.a. an ‘intrinsic good’. How does one argue for this kind of claim? Let me try the following. A cup of tea can be a good thing for you *in terms of your own*

⁶³ I won’t be doing any kind of technical analysis of the term ‘intrinsic value’ to begin with – i.e. before getting into the argument – because that is bound to make things complicated that are actually very simple. I will try to be precise in my claims and use of words, hoping that this will be enough to be able to convey my main points.

⁶⁴ In the literature, the standard formulation of ‘Kantian axiology’ (as a ‘dualistic axiology’ that has for that reason come to be known as ‘Kantian axiological dualism’) has it that the good-will (virtue) is *intrinsically* and *unconditionally* good, whereas happiness (well-being) is *intrinsically*, yet only *conditionally* good (Bader 2015, pp. 175-201). This ‘Kantian axiological dualism’ relates closely to Kant’s *dualistic human psychology* and the theory of ‘practical reason’ that takes that psychology for its basis: the psychological thesis being the idea that the human mind is driven by the two forces of *sensibility* (in the form of self-love, in which the agent pursues her own happiness for its own sake) and *reason* (in the form of a ‘sense of duty’), which in the context of practical philosophy becomes a matter of there being two fundamental practical principles (prudence and morality) that pull in potentially different directions but nevertheless interact in terms of *practical priority* – the priority of course being that morality always has to take precedence in virtue of the *categoricity* of the moral principle, which is why happiness (i.e. the agent’s own happiness) is deemed *intrinsically, yet only conditionally good*: we pursue it *for its own sake* – as a final end – but only *within the limits of morality*. To be clear, I am assuming this kind of Kantian dualism in my entire argument, but such is not the main subject of this section. It is really a much more basic point that I am trying to make in the main text. When (in the main text) I talk about Kant’s dualistic axiology what I am referring to is Kant’s recognition (which I am using to argue against Welfarism as purported moral axiology) of two independent and irreducible dimensions of value, two distinct and heterogenous kinds of valences, namely the *moral* and the *felicific*.

personal benefit, but only (and thus the point) *if it brings you happiness*, i.e. if you enjoy it directly or if your drinking it leads to some further increase in the balance of your pleasures and pains (by relieving an ongoing pain, for instance, or by improving your health and thus giving you a longer enjoyable life), whereas *that* enjoyment and *those* further pleasures would be *immediately* or *directly* good (i.e. beneficial) for you. It's not like, having enjoyed those pleasures, you would still need a further thing obtaining for the relevant benefit (the relevant 'good') to obtain. More pleasure would just mean more of the relevant benefit, and more of anything else (many more cups of tea, for instance) would mean nothing – in the relevant sense – unless it brings you pleasure.

It then seems perfectly clear that we can explain the 'value' (the felicific or benefit-related value) of things like cups of tea in terms of their instrumental contribution to the creation of happiness; more importantly still, what this basic analysis reveals is that we can explain the value of happiness in a self-referential way, simply by pointing to happiness as something that contains the source of its own value, as something that is *good in itself* in the relevant, benefit-related, sense – as that without which things in general couldn't possibly be of any value *in that sense*.

So far so good. But now take this analysis that the 'Welfarist' is offering and reflect on the thesis that *all value* (of all morally significant value, i.e. all value that is of direct interest for moral philosophy, and so presumably also 'moral value' as such) ultimately *reduces to felicific value* (i.e. to the kind of value that happiness has intrinsically and other things have only extrinsically). For 'Welfarists', remember, not only is it the case that happiness (well-being) is intrinsically good, it is also the case that *nothing else* is; not only is it the case that happiness can confer or impart value on other things, it is also the case that *nothing else* can. We have the internal value of happiness itself (as intrinsic value), and the value that happiness confers on other things by simply following from them as a consequence (as extrinsic value). That's it! That's all the 'value' that there is in the Welfarist picture: intrinsic felicific value, and extrinsic, but still always felicific, value.

What Welfarists are claiming, then, is that the kind of analysis (in reference to the intrinsic value of happiness) that we apply to cups of tea in our attempt to account for their *substantively beneficial character* can be used to fully account for the value (the morally significant value, and even, presumably, the 'moral value' as such) of all other things. Why

is the *physical health* of sentient beings a good thing? Because it normally sustains and enhances their happiness. Why is *death* normally a very bad thing? Because it extinguishes all prospects of happiness. Why is beauty important to the extent that it is important? Because of the rejoicing of those who stand in its presence. Why is *virtue itself* (moral activity) such a wonderful thing? Because, and only because, it puts us in the path of an ever happier world. And so on. Let me ask: is that the right story to tell in the context of our *moral axiology*? More importantly still: are we supposed to accept this story as an attempted ‘moral axiology’ at all? The answer to both questions is a resounding *no*, an answer that I give on the basis of the truth of the ‘Kantian thesis’.

But I cannot jump straight into defending the ‘Kantian thesis’ without first qualifying what its pretensions are, especially *vis-à-vis* the claims of Welfarism, some of which I argue (and Kant argues) are true (G 4:394):

A good will is good not because of what it effects, or accomplishes, not because of its fitness to attain some intended end, but good just by its willing, i.e. in itself; and, considered by itself, it is to be esteemed beyond compare much higher than anything that could ever be brought about by it in favour of some inclination, and indeed, if you will, the sum of all inclinations [**a.k.a. happiness**]. Even if by some particular disfavour of fate, or by the scanty endowment of a stepmotherly nature, this will should entire lack the capacity to carry through its purpose; if despite its greatest striving it should still accomplish nothing, an only the good will were to remain (not, of course, as a mere wish, but as the summoning of all means that are within our control); then, like a jewel, it would still shine by itself, as something that has its full worth in itself. **Usefulness** or **fruitlessness** can neither add anything to this worth, nor take anything away from it.⁶⁵

I take this to be Kant’s principal statement of what I am calling the ‘Kantian thesis’, and I will begin by telling you what I think Kant is doing here. Kant is advancing claims concerning the moral import of the *utility* and *disutility* of the good-will (its conduciveness or otherwise to happiness, presumably *universal happiness* since he is talking about the strivings of a morally good will), and these claims are (just be sure) all *negative* in nature.

⁶⁵ All emphases in bold are mine (as well as the square brackets).

What Kant is doing is in fact denying ‘utility’ (and so by implication also ‘happiness’, even if taken universally) any power to alter – i.e. to diminish or augment – the *moral worth* that attaches to the good-will (and therefore to human actions since human actions have moral worth only if done from a good-will), which Kant claims is *intrinsic* to it. This is of course the ‘Kantian thesis’. But denying ‘utility’ any *moral import* of the kind Kant is talking about in this context, i.e. any power to make human activity *moral* or *immoral*, *morally good* or *morally evil*, which certainly amounts to denying happiness (as that with respect to which the good will has utility) *intrinsic moral value*, is by no means the same as denying it *intrinsic value*. Thinking that these two things are the same is to think that ‘moral value’ is the only *kind* of ‘value’ that there is, an assumption that is unmotivated even as a reading of Kant (not to speak of the implausibility of an axiology that plays blind to the felicific as an independent dimension of value).

Again, it seems to me that saying (as Kant does) that the good-will can be *useful* or have *utility* – even in the context of denying this feature of the good-will any power to affect its *moral worth* – presupposes that there is ‘something’ *with respect to which* the good-will can be useful, namely – as we are assuming – *happiness* (everyone’s happiness). Now, ‘X being useful with respect Y’ seems to be the same as ‘X being a *means* to Y’; and something being a means to something else, or being useful with respect to that something, seems clearly to be one form of the *extrinsically good*: one form of that which is good in relation to *something else* that constitutes the *source* (the external source) of its value. Surely we have to presuppose that this thing with respect to which the good-will can be useful (*viz.* happiness) has a kind of value that doesn’t itself come from the good-will, for otherwise it would be senseless to even talk about the good-will as being *useful* (again, even in the context of denying this usefulness any *moral import*). Happiness, therefore, can be seen as that with respect to which the good-will can be useful, and the *intrinsicity* of its value is thereby presupposed. Its *moral value* (its power to alter the *moral quality* of human activity) is certainly being denied, but a certain notion of its intrinsic *non-moral* value is at the same time being affirmed. There is therefore a *qualified* ‘Welfarist’ thesis concerning the value of the human activity that holds true even in ‘Kantian axiology’: only to the extent that the good-will is in some way productive of happiness can we think of it as having a *beneficial* character. As far as felicific value is concerned, the good-will is very much like the cup of tea and the Sun: it

is no 'good' unless happiness follows from it (or is expected to follow from it, in case we are talking about 'expected value') as a consequence. This is the same positive story that we saw the Welfarist telling in relation to cups of tea and other things morality included, only now read into Kant's statements in his famous 'jewel passage'. But the main purpose of this passage is of course not to tell that story itself (which I am merely reading into it, and which is better told by Kant in other places), but rather to protect moral goodness *from it* by isolating it in thought (G 4:394):

Even if by some particular disfavour of fate [...] this will should entire lack the capacity to carry through its purpose; if despite its greatest striving it should still accomplish nothing, *an only the good will were to remain* (not, of course, as a mere wish, but as the summoning of all means that are within our control); then, like a jewel, it would still shine by itself, as something that has its full worth in itself. Usefulness or fruitlessness can neither add anything to this worth, nor take anything away from it.⁶⁶

I think we have to be with Kant on this point. The distinctive value of the good-will (its *moral worth*) has to be isolated from felicific considerations altogether by saying that happiness (and conduciveness thereto in any form) is *not* its source. However, if the question is whether happiness constitutes the source of the *felicific* value (the substantively beneficial character) of the good-will, the only answer can be yes, of course. It is as much as analytic (a.k.a. obvious, tautologous even) that happiness constitutes the ultimate source of the good-will's *felicific* value. But wait a second: isn't it equally obvious (tautologous even) that the good-will ('moral virtue') alone is of itself *morally good* and that the 'moral value of happiness' has to reduce thereto in like manner?⁶⁷ It is, and so I conclude that 'Welfarism' is false if taken as a purported *moral axiology*. Such is my main argument for the 'Kantian

⁶⁶ My emphasis.

⁶⁷ I construe the property of happiness (everyone's happiness) of being *extrinsically morally good* (in perhaps the only way I can make sense of it) as follows: happiness (everyone's happiness) acquires a moral character (a *moral necessity* attaching to it as an end of ours in particular) in virtue of its pursuit being commanded by the moral law, or if you will in virtue of its pursuit *from duty* being essentially constitutive of morality, in virtue of it being something that the *morally good* person distinctively cares about.

thesis', but I need to argue for it more thoroughly to avoid being taken as simply playing a game of words.

6.3.2 The Moral Side of Things

'Utilitarians' (thinking of those who consider themselves 'Consequentialists') might have plenty to say to defend themselves at this point. For one, it is by no means clear that they are necessarily committed to portraying happiness – even if taken universally – as being of itself a *moral good*, or as passing for the *moral* dimension of things. They may clarify their position by pointing out that their proposed 'moral axiology' (Welfarism) is not really a 'moral axiology' in the strictest sense of the term (i.e. an account of *moral goodness*, at least not initially). Welfarism, they could further clarify, has been designed to tell us about the '*non-moral* good' that morality essentially refers to as *its end* (the end – the non-moral good – the pursuit of which morality essentially involves), and they can tell various philosophical stories about what exactly morality's reference to everyone's happiness as an end amounts to. They could, for instance, decide to be 'Consequentialists', making *morality* a matter of – to put it simply – *conduciveness to everyone's happiness* (be it in actuality or in expectation, so there are *at least* two possible stories to be told here).

In this kind of analysis, the 'properly moral' side of things (e.g. the morality of actions) is not understood to lie in happiness itself; rather, morality is being identified with 'utility' understood as the property of *human actions* of being conducive to everyone's happiness (be it in actuality or in expectation). The word 'Utilitarianism' itself suggests that 'utility' has *something* to do with morality, and 'utility' is not the same as happiness (even if taken universally). Happiness doesn't equal *utility*, since utility is '*conduciveness to happiness*'. Bentham was explicit on this point (Bentham 1823/197, p.2):

By utility is meant that *property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness, (all this in the present case comes to the same thing) or (what*

comes again to the same thing) to prevent the happening of mischief, pain, evil, or unhappiness to the party whose interest is considered.⁶⁸

Utility doesn't equal happiness, then, and Consequentialists (i.e. Utilitarians who regard themselves as Consequentialists) are not saying that happiness obtaining (felicific outcomes considered just as such) is the same as morality or moral goodness obtaining. Instead, what they are saying is that morality equals (or in any case depends on) *utility* as the *conduciveness to everyone's happiness* of human conduct taken in some form.

Now, as between the options that these Consequentialists have at their disposal in this area, they do much better defining morality (the moral quality of actions) in relation to utility *in expectation* and not so in actuality, as actual felicific consequences are themselves beyond the control of intentional agents like us and morality is something that we rightly assume must be entirely within our control. Consequentialists can then begin to internalise morality and put it in human hands by arguing as follows: in her intentional activity, the moral agent is *intending to put something out there* (to promote everyone's happiness), but once she has acted, once it is out there, it is no longer up to her, and therefore no part of her morality. Morality is this '*intending to put something out there*', as opposed to something out there actually occurring. I submit that these thoughts are all good thoughts, thoughts that actually say something about the nature of morality. However, I argue that they have to be embedded in the right sort of moral framework, which, I also argue, is *not* a 'Consequentialist' framework of any sort.

The 'Consequentialist' framework that Utilitarians commonly endorse is a fit for *Welfarism*, but *Welfarism* (which I submit is the starting point for most Utilitarians, because they get to Consequentialism by first thinking that happiness is that which ultimately matters and then realising that what ultimately matters about human conduct is the impact that it has on everyone's happiness) is false, at least as a *moral axiology* in the proper sense of the term (i.e. as an account of moral goodness and moral evil, of morality and immorality as distinctively moral valences). We see the problem in the quote by Bentham that we have

⁶⁸ My emphasis.

been discussing in relation to the meaning of utility [although now with a different emphasis to indicate the new point] (Bentham, 1823/1907, p.2):

By utility is meant that property in any object, whereby it tends to produce *benefit, advantage, pleasure, good, or happiness, (all this in the present case comes to the same thing)* or (what comes again to the same thing) to prevent the happening of *mischief, pain, evil, or unhappiness* to the party whose interest is considered.

Is Bentham justified in equating happiness (pleasure, advantage, benefit) with ‘good’ or ‘goodness’ itself? If you bracket *moral goodness* (the kind of goodness we ascribe to saints, for instance), perhaps. But we don’t want to bracket *moral goodness* because moral goodness (that which constitutes moral goodness, namely the human will in the realisation of its full moral relationality) is the very subject matter of moral philosophy. Is Bentham justified in equating unhappiness (pain, mischief, disbenefit) with ‘evil’? Again, bracketing *moral evil* (the kind of evil that we ascribe to scoundrels, for instance), maybe the only other negativity that remains which one may want to call ‘evil’ is pain itself, I agree. But obviously we cannot just bracket *moral evil* in ethics.

Now, it is true that Welfarists (Utilitarians acting as Welfarists but also trying to defend their Consequentialist commitments) may at this point *insist* (on a point that I noted at the top of this section) that they are not actually bracketing these things (moral goodness and moral evil) and forgetting about them but instead offering *their own account of them*. Having agreed on felicific goodness being a *non-moral* kind of goodness, and having put ‘Consequentialism’ in place as the normative framework that is meant to account for the ‘moral side of things’ (thus giving us a certain notion of the ‘morality’ or – as it is more often called the ‘moral rightness’ – of actions), they could define ‘moral goodness’ as the fundamental disposition of the *moral subject* to reliably do ‘the right thing’, which as such is assumed to be felicifically optimific in expectation. This may sound like the kind of thing that we distinctively ascribe to saints: the disposition to always do the right thing and to thereby maximise the chances of happiness for all sentient beings. But then of course we can imagine someone who is reliably and systematically disposed to do the right thing (and therefore also the felicifically optimific thing in expectation) but only *out of fear of being*

doomed, of being placed in Hell by a God who knows her every movement and expects her always to do the right thing. But I wonder: is she really being a *morally good* person? Are her apparently diligent actions *morally good* (genuinely moral) actions? I don't see how they could be, and this thought alone shows that there has to be a more basic conception of 'moral goodness' that escapes 'Consequentialist' analysis. Only when we think of the person as acting *from a distinctively moral motive* (i.e. *from duty*, or *in practical love* if you will, or even *for everyone's happiness*, or all of the above together as it is the case in the final account) do we start to see any *moral goodness* in the action.

I can accept that 'Consequentialists' are free to point to the *moral subject* and her own volition (her intentions, what she intends to do) in their attempt to situate the morality of actions, thus initially deflecting the worry that they may be completely blind to the 'human' ('properly moral') side of things; but I don't think their project can ever succeed, as 'Consequentialism' is itself the result of avoiding reflection on moral valences (morality and immorality) properly so called, which we attach to the causality of the human will *as such*, as the special kind causality that it is in its different modes. Again, 'Consequentialists' may try to give humanity its due by pointing to human activity (even to our most fundamental disposition or maxim) as that which constitutes the 'properly moral' side of things; but the second they start judging this activity on the basis of 'Consequentialist' standards (which as such can be applied to assess the axiological – always non-moral – merits of all kinds of things, as Consequentialists themselves have acknowledged in ideas such as *Global Consequentialism*), they immediately undo the favour. Consequentialist evaluation is *non-moral evaluation*, and it remains the same even when applied to those things that actually of themselves constitute 'the moral side of things' (human maxims, motives, intentions and actions).

Nothing against Utilitarianism itself has been said here, of course, only against the formulation of the view that has become standard in academia. Remember that in the final account Utilitarianism is to be construed as an 'Agapist' (as opposed to 'Consequentialist') doctrine depicting morality as a matter of *practical love*, of the agent being *benevolent* (in a very precise sense that has to be described in detail, which is one of the things that I attempt to do in **Part II**) to all sentient beings. But depicting morality as a matter of exercising practical love (which as such involves acting *for everyone's happiness*) is to depict morality as

irreducibly a matter of *attitude* on the part of individual moral agents, as a matter of these agents acting on good motives (principles, maxims) *as opposed to* producing or even ‘promising’ to produce (in the ‘expected value’ sense as applied to their conduct) the best felicific consequences.⁶⁹ Morality is therefore neither in happiness itself (even if taken universally) nor in the conduciveness thereto of human activity (i.e. in *utility*, be it in actuality or even in expectation), as it lies only in *practical love for all sentient beings*. This, according to Utilitarianism as it should be.

Having concluded my basic argument against the truth of both Welfarism and Consequentialism (and to bring this **Part I** finally to an end), I will end by giving you some further reflections on *moral goodness*.

6.3.3 On Moral Goodness

My general message so far has been that ‘moral goodness’ is as much as a *misnomer* when predicated of either happiness as such (intrinsic felicific goodness) or conduciveness to happiness (utility) as such (extrinsic felicific goodness). Even the most general reliability of the human will (considered in the most basic acts that it performs) to do that which ‘promises’ the best felicific results in an indefinitely ongoing future cannot as such be called *morally good*. Which *of course* is not to say that the morally good person is indifferent to the promotion of happiness in the world, but nevertheless. ‘Moral goodness’ and ‘moral evil’ are the basic terms we use in our practice of *moral* assessment, which is always a *judgement of reason* targeting human conduct regarded as *free*, i.e. as being capable of guidance through

⁶⁹ If there is a *promise* that is of itself part of morality in relation to the happiness of subjects in a fundamental kind of way, it is the promise (in the literal sense of the term) that the human being address to all sentient beings (if only in thought) to be by their side to protect them for their own sake, to do her job as a guardian in a moral community that includes all of them as citizens with beneficiary status. The morally good person acts on this (as I would like to call it) *agapist* promise, and this manner of acting is essentially constitutive of *morality*. Then of course, we can assume, comes also the ‘Consequentialist promise’ (the sheer fact that – to put it simply – the probability of a happy world coming to pass has *thereby* – as a result of the morally good activity of the subject – been maximised); but this ‘promise’ (which goes by the name of ‘expected utility’) is not of itself part of morality.

the influence of reason in the form of a moral command addressing the human will *directly* (responsiveness to the *moral law*). Only obedience to the moral law in the privacy of human wills (only this obedience taken *as such*) can make human conduct in any proper sense *morally good*, a.k.a. genuinely moral; it takes obedience to the moral law in the privacy of human wills for things to improve *morally*, and all felicific properties of human action are on this account rendered morally immaterial (i.e. in the sense that is directly implicated here: the incapacity of felicific properties in general to alter the *moral quality* of human activity).

Now, people normally use the basic axiological terms ‘good’ and ‘bad’ in reference to the felicific as such. They say things like: I had a *good* time, that was a *good* concert, the pain was *horrible*, something *very bad* just happened, and so on. In my view, they have every right to do so, and I have been portraying Kant as agreeing with this judgement in virtue of his general axiological stance allowing for the felicific as an independent (irreducible) dimension of ‘value’. However, there are a few statements in Kant that *seem* to speak to the contrary, and we would do well to examine those statements, including (CPrR 5:60): “Thus good or evil is, strictly speaking, referred to actions, not to the person’s state of feeling.” The most basic of axiological concepts – Kant tells us in this line – are ones that can only be predicated of human actions as such (regarded as ‘deeds of freedom’, would be the assumption here), and he seems to be making a closely related point when – only a few pages below – he states that the moral law is conceptually prior to the concepts of ‘good’ and ‘evil’ (CPrR 5:63): “the concept of good and evil must not be determined before the moral law (for which, as it would seem, this concept would have to be made the basis) but only (as was done here) after it and by means of it.” These two lines are taken from Kant’s *Second Critique*, where he also explicitly distinguishes between the concepts of ‘*das Gute*’ or ‘the good’, on the one hand, and ‘*das Wohl*’ or ‘well-being’ on the other (CPrR 5:59–60), thus giving to understand that well-being doesn’t of itself deserve the status of a ‘good’. And so I ask: What are we supposed to do with these statements by Kant apparently denying axiological status to – talk of ‘goodness’ and ‘badness’ in relation to – the felicific as such? Here is my suggestion (i.e. in relation to our reading of Kant, which is only secondary to the independent points I am trying to make concerning ‘moral goodness’).

If we read Kant's claims (those seemingly denying axiological status to felicitic items as such) as expressing *unqualified* propositions, then a certain kind of 'Kantian axiological monism' would seem to follow unavoidably where the good-will is not only intrinsically good, but as much as *the only* 'good' properly so called. We could on this basis read Kant as engaging in the revisionist project of trying to reform our common axiological discourse, i.e. given that people (everywhere as far as I know) do in fact use axiological concepts such as 'good' and 'bad' ('evil' even) in relation to the felicitic as such. But I think it is possible to construe Kant's apparently problematic statements in such a way as to avoid putting him in that position: Kant's statements need not be read as unqualified propositions, as they can in fact be qualified in various ways to convey the important nuance that they mean to convey (as I will try to show after a small reflection on why 'Kantian axiology' has to be construed as dualistic at certain fundamental levels).

Thus, I submit that the claims developed so far in this section (e.g. the claims that I managed to read into Kant's 'jewel passage') speak strongly in favour of a reading of Kant as giving us the distinction between 'moral goodness' and 'felicitic goodness' as two perfectly distinguishable, heterogenous kinds of 'goodness'. But there is explicit textual support as well in favour of that reading. In the *Second Critique* (the place where Kant makes some of his apparently problematic statements) he identifies moral virtue with 'the *worth* of the person', and well-being with 'the *worth* of the person's condition' (CPrR 5:60), meaning that there are two distinct kinds of *worth* which persons (human beings in the context of Kant's discussion) can realise in their lives. This is of course the same distinction as that between someone's being a *morally good* person (moral goodness) and someone's 'doing well' or benefiting out of life (well-being / felicitic goodness / happiness). Not only are there two heterogenous 'goods' which persons (human beings) can realise: these two 'goods' (moral virtue and happiness) actually come together to constitute what Kant calls the "highest *good* in a person" (CPrR 5:110), or 'the good of human beings' as we might also call it.⁷⁰ Kant conceives of this 'highest good' as the realisation of all that is intrinsically (irreducibly) good in the life of human beings, and consequently he also refers to it as the

⁷⁰ To be distinguished from *the human good* properly so called, the *distinctively human good*, which of course has to be *moral virtue* alone.

‘complete good’ (CPrR 5:110). Happiness may be a *non-moral* good, but it is a good nevertheless, a good that is as much as constitutive of (essential to) the realisation of human beings *as persons* (i.e. as both rational *and* sensitive beings).

It is worth noting that it is one of Kant’s major goals in this area of his philosophy to arbitrate between the Stoics (the champions of virtue) and the Epicureans (the champions of pleasure). Kant rejects ‘Epicurean’ philosophy because he takes it to misportray morality as the doctrine of how to *make oneself happy* (e.g. CPrR 5:112, 5:126), but he also rejects ‘Stoic’ philosophy because – among other things – it misportrays both virtue and pleasure: (i) virtue as being ‘its own reward’⁷¹ (which he strictly speaking denies),⁷² and (ii) pleasure as an ‘extrinsic good’ or an ‘indifferent’ (if only a ‘preferred indifferent’) from the perspective of one’s own happiness or well-being, which it plainly isn’t according to Kant (whom I read as an unqualified hedonist about well-being as such, which *of course* is not to deny that the pursuit of one’s well-being in self-love is always conditioned by moral duty in Kant’s theory of practical reason).⁷³ The synthesis of the two views that Kant proposes (which he associates with ‘Christian philosophy’) has it that ethics (the doctrine of moral virtue) is not to be understood as a doctrine of how to ‘make oneself happy’, but instead as a doctrine of how to make oneself *worthy of happiness* (CPrR 5:130). But if virtue has the sense of *worthiness of happiness* or *worthiness of reward*, then it cannot be

⁷¹ There are other ways in which Kant sees the ‘Stoics’ as misportraying virtue (CPrR 5:126–127) and the ‘Christian’ doctrine as superior (i.e. truer), but I cannot afford going into the details of this right now (see **Part III**).

⁷² To be fair to Kant, he does try to accommodate the idea that ‘virtue is its own reward’ in the idea of a “moral pleasure” (MM 6:391). But a ‘moral pleasure’ is a pleasure, of course, and consequently there is no argument here against ‘hedonism about well-being’ (not even against *unqualified hedonism*, which doesn’t care about the nature or the provenance of the pleasure for the sake of counting it towards well-being just as such).

⁷³ It is to this effect that Kant replies to the Stoic’s cry [in italics for emphasis]: “*Pain, however you torment me I will still never admit that you are something evil*” (CPrR 5:60). Kant concedes that the Stoic is right in thinking of pain as something that doesn’t in any way diminish the worth (the moral worth) of his person, but that nevertheless his cry “betrayed” the fact that “he felt that the pain was an ill” (CPrR 5:60). In other words, Kant would have liked the Stoic to admit that pain (no matter what he makes of it) is very much something that detracts from the *worth his condition* (i.e. from his well-being, which always comes down to hedonic states and as such is *essentially* the same in kind as that of any other sentient creature).

thought of as itself having the sense of *reward*. Virtue is not (nor does it of itself contain) its own reward. Only *happiness* understood as something distinct from virtue (*viz.* to put it simply, pleasure) has the proper sense of *reward*. The Epicurean view has then been saved in the latter idea,⁷⁴ and the Stoic view has been saved in the idea of the purity (the disinterested nature, or more precisely the *purely rational* nature) of morality. However, it's only the positive insights of both views that have been saved, and the reductive (monistic) pretensions of both have been rejected in favour of a *dualism* that has begun to take form in the ideas of the *highest good* as the sum of two heterogenous 'goods' (moral virtue + happiness) and the related idea of *remunerative / punitive justice* (happiness *in proportion* to virtue).⁷⁵

Kant's insistence on a certain idea of *justice* (happiness in proportion to virtue) and the related idea of the *highest good* (moral virtue + happiness) are usually taken to be essential to his view, and therefore also to any credible reconstruction of Kantian philosophy. I don't mean to deny these claims here, but I am for the moment taking the ideas in question as 'problematic' (as ideas yet to be discussed), and the reason I am appealing to them in this context is only to make a substantive point of my own: I have been using them as *evidence* that Kant is (in the context of our present discussion) a *dualist* insofar as he grants the intrinsic value of both virtue (morally good willing) and happiness or well-being (pleasure). To be perfectly clear, the point is not exactly to 'prove' that Kant is a dualist in this context, or even to defend some reading of Kant on purely exegetical grounds, but simply to seize the opportunity afforded by the themes being discussed (i.e. the opportunity to use my own

⁷⁴ This is of course Kant's *hedonism* about well-being (CPvR 5:60): "Well-being or ill-being always signifies only a reference to our state of agreeableness or disagreeableness, of gratification or pain, and if we desire or avoid an object on this account we do so only insofar as it is referred to our sensibility and to the feeling of pleasure or displeasure it causes."

⁷⁵ Please don't take me to be arguing here for the validity of these two ideas, since it's mostly in the **Part III (Conclusion)** that I get to discuss them in any detail. To anticipate, I will for the most part be avoiding pronouncements on the validity and significance of Kant's idea of remunerative / punitive justice in particular, although I will be explaining the reason for this decision towards the end of that concluding section.

understanding of Kant's philosophy) to convey certain specific points that I need to convey concerning *moral goodness*.

Let's go back to Kant's idea of *remunerative justice* for a second. The reason why happiness can be seen as having the sense of *reward* is because it has the more fundamental (more indigenous to it) sense of *benefit* (i.e. of 'substantive benefit' to use my proposed terminology); in similar manner, the reason why moral virtue can be seen as having the sense of *worthiness of reward* is because it has the more fundamental (more indigenous to it) sense of fulfilment of a *mission* or *task* attaching to our existence as human beings (human dignity).⁷⁶ You fulfil your *mission*, you stand to receive a *benefit*, is the idea. Now, the conclusion really seems to follow from this simple analysis that moral virtue and happiness have fundamentally different *senses*; each is 'good' in its distinctive (indigenous) kind of way. As to the good-will or moral virtue (for facility: *acting from duty*), the goodness that we distinctively attribute to it (moral goodness) is nothing but the recognition of it counting as the state of realisation of the human being as such, i.e. as realising that which makes a human being (in her own distinctive intentional activity) a *good* human being, as that which makes actions (as pertaining to the distinctive activity of human beings) *good* actions. Any ascription of moral goodness (to persons or their actions) is thus the recognition that there is a *human being fulfilling her moral (distinctively human) mission in the world*.⁷⁷

⁷⁶ This point was not needed in the main text, but one could certainly argue as follows: Kant claims that happiness and nothing other than happiness is *deserved* (i.e. intrinsically, since of course one can deserve many things in virtue of deserving happiness) by people to the extent that they act virtuously. But how is this possible? It seems perfectly clear that one of the conditions of the possibility of this 'deserving of happiness' is the idea that there is something to happiness, and to nothing other than happiness, that makes it uniquely *deservable*, and what could this be apart from an internal value attaching to happiness (having the sense of *benefit*) that is prior to its actually being deserved? Happiness has 'intrinsic deservability', we might say. I mean, in Kant's view, not that this is the only or even the best way to speak of the intrinsic value of happiness. Given the point of my argument, my emphasis in relation to happiness is one concerning the indigenous sense of *substantive benefit* that it has.

⁷⁷ Many questions arise at this point in relation to attributions of moral goodness, especially in relation to categories that have been introduced by Utilitarians: e.g. the distinction – which has to be understood in a specific way – between the 'morality of the act' and the 'morality of the agent' (e.g. Mill 1861/2003, p.196). I address the issue (although only in passing) when in **Part II** (section 7.1) I get to *propose* that this particular

There is one last thing to be done, which I left in the air when I began listing general reasons for keeping ‘Kantian axiology’ as I have portrayed it in this work (as – among other things – preserving the axiological status of the felicific as such). What we are yet to do, of course, is address the question what to make of Kant’s seemingly problematic statements concerning the axiological status of the felicific as such. Let me bring these statements back for consideration: first we have the claim that “good or evil is, strictly speaking, referred to actions, not to the person’s state of feeling” (CPrR 5:60); which is compounded with the claim that there is ‘*das Gute*’ or ‘the good’, on the one hand, and ‘*das Wohl*’ or ‘well-being’ on the other (CPrR 5:59–60). But surely we can understand Kant’s claims in this context as referring to *moral goodness* in particular as that which (together with *moral evil*) is essentially referred to *human actions* (as deeds of freedom) and is always to be distinguished from *well-being*, and we seem to be equally free to qualify ‘goodness’ in reference to the latter: *felicific goodness*.

Kant’s other apparently problematic statements (i.e. problematic from the perspective of trying to preserve the axiological status of the felicific as such in our ordinary discourse) add more dimensions to our discussion. I cannot fully address the issues here, but I can make a few relevant points (CPrR 5:63): “the concept of good and evil must not be determined before the moral law (for which, as it would seem, this concept would have to be made the basis) but only (as was done here) after it and by means of it.” Again: if qualified as referring to ‘moral goodness’ and ‘moral evil’ in particular, I agree with this claim. The concept of *moral goodness* cannot be defined without the concept of the moral

distinction that Utilitarians have used better be mapped onto Kant’s basic distinction between ‘morality’ and ‘legality’. This move admittedly loses the distinction itself between the ‘morality of the act’ and the ‘morality of the agent’ as Utilitarians have intended it, as in the Kantian picture there is only *one morality* (having its ultimate source in the fundamental maxims of the human will) that is equally predicated of *both* persons and their actions (the distinction being more or less artificial depending on the context). The promise, of course, is that this move will not damage but strengthen the credibility Utilitarianism as an ethical theory (since now we don’t need to call *actions* ‘moral’ – but perhaps only ‘right’ or ‘legal’ to the extent that they are, knowing that none of this will ever pass for ‘the *morality* of actions’ – that are done from merely self-serving motives or, worse still, from a morally evil maxim).

law, since moral goodness *just is* (in the sense of being essentially constituted by) *obedience to the moral law* or (as in common parlance) *acting from duty*.

The concept of what is ‘good’ in a wider practical-rational sense is also one that has to refer to the *moral law* as giving us the *unconditioned condition* (the supreme limiting condition) in our practical life, as Kant had emphasised already in the *Groundwork* in the claim that “nothing has any worth other than that which the law determines for it.” (G 4:436) If, for instance, there is a kind of pleasure the pursuit of which is inconsistent with our morality, then surely we shouldn’t be calling it ‘good’ in any sense entailing that we have reason (justification) to pursue it. It is in this context that one is most tempted by the revisionist project of our common axiological discourse in general that we might read Kant as trying to propose: in relation to the eating of flesh, for instance (and you can imagine more examples), I would like to say: pleasant, perhaps, but *no good*. But I think I have every right to say this (i.e. to say it – and so in a way join Kant’s project – without giving up the points that I have been making) because this way of speaking makes perfect sense *in the relevant context of utterance*. In the context of giving justifications for conduct that may be challenged on moral grounds, it is just unacceptable to point to the pleasure one would derive from such conduct.

The conclusion that I draw from this basic reflection is the following: so long as terms are properly qualified, and so long as the use of terms be adequate given the context of utterance, I claim that there is no philosophical problem with calling pleasure ‘good’ and pain ‘bad’ (‘evil’ even).⁷⁸ There is a problem with calling certain pleasures ‘good’ in

⁷⁸ The flipside here is that very little is gained philosophically speaking by being able to call pleasure ‘good’ and pain ‘evil’ (i.e. to call them by these terms considered just as such). Something that is gained (I mean, something that we won’t lose) is the right to talk about the happiness and unhappiness of animals in a certain way. The pain of an animal is a bad thing even in isolation from moral or rational considerations (in isolation from *oughts* of any kind), meaning of course that we don’t have to apologise for calling it ‘bad’ and instead call it only ‘painful’ or ‘disagreeable’ (just like we don’t have to apologise for calling the good-will ‘good’ instead of just ‘moral’ or ‘noble’, which I can imagine someone insisting that we do). After all, what I have been doing here (in so far as I have been arguing for preserving the axiological status of the felicitic as such) is simply recognising the existence of an independent dimension of ‘value’ that merits the existence of a sub-field in philosophy which of course *relates* closely to other fields: the theory of well-being, in the context of which well-being itself counts as ‘the good’. Having said this, I will insist on the idea that very little has been

certain (perhaps most) contexts of utterance, but that is something that can easily be solved in reference to the moral theory and the more general theory of practical reason (where morality figures as the unconditioned) that define the framework for our constructive dialogue as human beings.

If there is a part of ‘moral goodness’ that I would like to protect from any kind of qualification is not ‘goodness’, then, but *moral*, and so I want to address philosophers working on these issues (Utilitarians in particular): please qualify your use of axiological terms to always be very clear in what you mean, but never give me the term ‘moral’ in reference to the felicitic *as such*. Philosophers are deluded who think that calling pleasure ‘good’ and pain ‘evil’ (even in conjunction with ‘Consequentialism’) will ever pass for a *moral axiology*. Moral goodness (perhaps the main *explanandum* in moral theory) belongs in exclusivity to the realisation of the human being as such in the instantiation of a morally good will (human dignity), and in recognition of this fact we have to reject ‘Standard Utilitarianism’ and join the project of ‘Kantian ethics’. But I argue that the ‘standard utilitarian formula’ (‘Consequentialism’ + ‘Welfarism’) is far from the correct formulation of *Utilitarianism* (which I argue has to be construed as an *agapist* doctrine that is perfectly consistent with the basic Kantian framework I have begun to put in place and will continue to develop in **Part II**). Let me now give you my own formulation of *Utilitarianism* and my defence of *the dignity of all animals*.

gained philosophically speaking. For instance, if one wants to call happiness (e.g. the happiness of non-human animals) ‘good’ in a *moral* sense (as something that we *ought morally* to pursue, for instance) one needs to produce an independent theory accounting for the relevant *moral ought* (which we now understand has to be a theory of moral goodness as a *sui generis* kind of goodness belonging in a theory of human dignity in the Kantian sense of the term). One cannot get there by simply pointing to the felicitic and insisting on its independent axiological status (e.g. by pointing to pain as something that is ‘just bad’), nor by simply pointing to the felicitic tendencies of human activity. We need a theory of human moral relationality putting us in the right moral relation to sentient beings; hence my main objective in **Part II**.

Part II: Animal Dignity: A Defence of Utilitarianism

Part I of this work was devoted to developing and defending the basic claims that constitute the Kantian doctrine of *human dignity*. In parallel, the argument was also made that we have to reject the ‘standard utilitarian formula’ (Consequentialism + Welfarism) in favour of a Kantian conception of *morality* making it depend entirely on the intrinsic quality of the human will. Felicific consequences are immaterial to the *morality* of human actions because – as we normally say – it’s the *intention* (the intrinsic quality of the individual’s volition, the *maxim of the will* ultimately) that counts!

Part I thus also left us with the thought that morality is necessarily *logo-centric* (or, to use the somewhat less accurate but more popular term, *anthropocentric*) in a very central and crucial sense. Moral standards are the very standards our humanity (our morally rational nature), from which it follows that the ground of all moral realities is to be located in human life as such, i.e. in the *existence* of human (morally rational) beings acting in the world. If it weren’t for the presence of human beings, if it weren’t because there are beings whose conduct stands under (self-represented) moral standards, beings who by their own nature carry the burden of moral duty, *moral goodness* and *moral evil* would be completely out of the question, and *moral relations* in general would be strictly speaking impossible. No wronging relations could exist in such a world, for instance, since there would be no potential wrong-doers. Human beings are in this way the protagonists of the show, the very ground and locus of all things moral.

Now, it has never been my claim that human beings are superior to other animals in every morally significant respect, and the main goal of this **Part II** is in fact to argue for the notion of *animal dignity*, to argue that there is a certain morally significant domain calling for the use of this notion.⁷⁹ There is a lot to be said about what the notion of ‘animal

⁷⁹ I will later in the argument be questioning the identity between ‘animality’ and ‘sensitivity’ (the capacity for happiness + concomitant agency) on the basis of the thought that ‘animal’ equals ‘person’ equals ‘subject’ as one and the same general philosophical category, thus in principle allowing for the existence of ‘purely rational (i.e. non-sensitive) animals’. To prevent misunderstandings, let me then warn you that when – in the

dignity' amounts to, but the basic point is actually very simple: *animals in general* (all animals *qua* animals, i.e. all sentient beings *qua* sentient, all beings capable of happiness *as so capable*) have a certain *moral standing* – whose nature we need to elucidate – raising them categorically above the non-animal (non-sentient) world. The claim here is not that animals (sentient beings) occupy a middle place between inanimate nature and rational beings (which of course they do in some sense), but rather that there is a *special place* in morality for sentient beings *qua sentient*, a place that not even the Highest Being (in case he were incapable of pleasure and pain) could occupy. This should be enough for the time being, but you can expect me to develop the claim further as the argument goes.

Importantly to be noted is that the title of this **Part II** makes the connection explicit between the project of making sense of the notion of 'animal dignity' (the dignity of sentient beings) and the project of defending *Utilitarianism* as an ethical theory. And that is because – as I see it – the notion of animal dignity holds (as a moral category) *because* Utilitarianism is true. Since the 'standard utilitarian formula' has been dismissed as inadequate, and given that a basic Kantian conception of morality and related ideas have already been established (if only for the purposes of this project), the argument ahead will have to include at least two things: (i) an explanation of what Utilitarianism actually amounts to, and (ii) an explanation of how Utilitarianism (thus understood) fits into the basic ethical framework put in place by Kantian ethics.

But it's not only the nature of Utilitarianism that is problematic (unclear) at this point, since it is a similar story with Kantian ethics: we have fixed some ideas and concepts in the name of this view, but all of that is really nothing more than a very basic starting point. The promise is that in the final analysis – and very much against common conceptions of how things go in this area – it will be possible to say that animal dignity (the dignity of sentient beings) holds because Kantian ethics is true, because there is no real conflict between Kantian ethics and Utilitarianism, or perhaps even more to the point:

argument that follows – I speak of 'animal dignity' (or 'distinctively animal dignity') what I mean is 'the dignity of sentient beings *as such*', 'the dignity of individuals or subjects *as capable of happiness*'. I will be constantly qualifying my claims anyway, so that confusions shouldn't arise.

because the concepts of morality and (specifically) human dignity necessarily entail the dignity of all sentient beings.

As a final warning, let me note that I have decided to follow the same general pattern of argumentation that we saw in **Part I**, this time giving the first word to Kant and the dominant Kantian doctrine (which rejects the notions of ‘animal dignity’ and ‘respect for animals’) as my initial antagonist.

1. Kant and the Rejection of ‘Animal Dignity’ and ‘Respect for Animals’

What is Kant’s story in the context of ‘animal ethics’, then? The first thought that Kant gives us is that there are two – and only two – categories to which individuals can belong in terms of their basic *moral standing*. Namely: individuals are either *persons* (beings with ‘dignity’) or they are *things* (beings, not with dignity, but only with ‘price’) (G 4:428). This is of course one of the basic distinctions in Kant’s philosophy (as well as common discourse) that this work embraces: after all, it is one of my main goals in this work as a whole – if not *the* main goal – to offer an account of *the dignity of persons*. Now, Kant’s view becomes problematic when he gives us a second thought, namely that humanity (or lack thereof) is *the only* criterion for membership in each of the two basic categories just introduced. The thesis, in other words, that *only human beings* count as ‘persons’.

Again, if all and only persons have dignity, and if – as Kant argues – all and only human beings count as ‘persons’, then it just follows that non-human animals in general count as ‘non-persons’, i.e. they count as *things*, and entities that belong in the category of ‘things’ have been stipulated to have ‘price’ and ‘price’ only (‘relative worth merely as means’), no ‘dignity’ whatsoever (G 4:428). Looking at Kant’s fundamental claims here, one can only conclude that the notions of *animal dignity* and *respect for animals* have no place in his philosophy. But what exactly does it mean (morally speaking) to deny animals ‘dignity’ and declare them ‘things’ alongside inanimate objects? Let me quote a couple of lines by Kant which may give us a good initial sense of what he may have in mind:

The fact that the human being can have the “I” in his representations raises him infinitely above all other living beings on earth. Because of this he is a person, and by virtue of the

unity of consciousness through all changes that happen to him, one and the same person – i.e., through rank and dignity an entirely different being from *things*, such as irrational animals, *with which one can do as one likes*.⁸⁰ (*Anth* 7:127)

When [man] first said to the sheep, “the pelt which you wear was given to you by nature not for your own use, but for mine” and took it from the sheep to wear it himself, he became aware of a *prerogative which, by his nature, he enjoyed over all the animals*; and he now no longer regarded them as fellow creatures, but as *means and instruments to be used at will for the attainment of whatever ends he pleased*.⁸¹ (*CBHH* 8:114)

If in the final account it is shown that Kant’s fundamental principles (properly understood) are consistent with the notions of ‘animal dignity’ and ‘respect for non-human animals as persons’, it will be very much against Kant’s own judgement. To be fair, Kant has much more to say about non-human animals than I have registered so far, including some partially redeeming claims that are potentially in tension with the ones we just read. I will be discussing those claims in due course. For now, however, that is what we get: a very negative – almost mocking – take on the moral standing of non-human animals.

The two lines by Kant I just quoted manage to express two main philosophical ideas: (i) that human beings are *higher in rank* with respect to the other animals, and radically so, and (ii) that human beings ‘naturally’ conceive of themselves in that way. In a certain sense, therefore, these are statements of *human dignity* in the substantive or non-reductive sense of the term: statements affirming the elevation (in moral standing) of human beings *as such* with respect to the non-human world (non-human animals in particular), an elevation which we ‘naturally’ (if only from our perspective as moral beings) see ourselves as having. We have what is normally called a ‘sense of dignity’: a sense of our *unique importance* as human (morally rational) beings.

However, it’s not like Kant is here – in these two lines – giving us those ideas of human moral superiority *themselves*: he is not telling us about how we are – and consider ourselves – singularly important in virtue of the *moral burden* that we alone bear. Instead,

⁸⁰ The second emphasis is mine.

⁸¹ Both emphases are mine.

what Kant is saying is that human beings enjoy a certain *prerogative* over non-human animals that is part and parcel of their dignity: as the superior beings that they are, human beings are *entitled* to do with non-human animals *as they wish* (as the first passage suggests), to take their skin and wear it themselves for instance (as the second passage suggests).⁸²

Don't get me wrong: Kant does have in mind our moral uniqueness when he makes his claims in this context. We can take it for granted that when he speaks of the "I" that is assumed in our representations – which raises us *as persons* "above all other living beings on earth" (*Anth* 127) – he actually means the 'moral I', i.e. the 'moral self', of which we become aware in our judgements, our practical activity and our very way of seeing the world.⁸³ It is this 'moral I' that – as Kant sees it – makes us so fantastically 'important' that a license to use non-human animals as resources (in the ordinary sense of 'supplies') obtains:⁸⁴ a recognition of our moral nature – Kant thinks – raises us so high in 'rank' and 'importance' in our own eyes that we now look at the other animals and just fail to see them as our fellow creatures. And so the thought goes: when I encounter a human being, I cannot say to her that her skin and flesh belong to me: I cannot say to her that she is alive to grow and fortify skin so that I wear it, or to keep meat fresh so that I eat it, and this simply because I recognise her as a fellow human being. If, on the contrary, it is a non-

⁸² When Kant says that we can treat animals 'as we wish', he doesn't mean to say that we can treat them as if there were no moral duties in place, which is how one may naturally (and understandably) read the claim. What Kant means, rather, I think, is that we have a prerogative to decide on the *general fate* of animals, i.e. to use them for food or clothing or work-force or companionship, etc. (thus treating them 'as we wish' in this general discretionary sense), but then of course we have a duty to 'treat them well' *whatever purpose we give them* (e.g. trying not to inflict 'unnecessary' suffering and death).

⁸³ Kant cannot possibly mean that only human beings have an *I*, since all animals (i.e. insofar as they are animals) clearly have one: many of the non-humans (the majority, if not all, of those whom we systematically exploit) are experiential subjects (obviously) endowed with various levels of self-awareness, complex emotional capacities, volitions, various kinds of intelligence and perhaps even rationality of some sort. Only human beings, however, have a *moral I* – only human beings have an *I* that stands in relation to self-represented *moral* standards. Only human beings are *morally* rational.

⁸⁴ One can describe Kant's general taken on the moral standing of animals is that they are 'resources', which I argue he means in the ordinary sense of 'supplies'. However, it is worth noting that perhaps the most important use we can make of animals (still regarded as supplies) is as equipment for our moral training, which we do by simply being *kind* or *benevolent* to them and caring for their happiness.

human animal that I encounter (a pig, say), now I am at liberty to say exactly those things that were out of place in relation to my fellow human. That skin which you wear is mine, give it to me! I can say. That ‘meat’ covering your bones is also mine, give it to me! I can say.⁸⁵ To speak allegorically, it’s as if God’s act of appointing human beings their moral mission (a great thing on its own) were necessarily bound up with the act of addressing the animals and the rest of the world *to human beings* (under the rubric of ‘God’s special love for human beings’) as a present or as a gift. Not only have we been given all ‘things’ for our own ends and benefit to the exclusion of the ends and benefit of other animals; these animals are *themselves* part of the ‘gift’ bestowed: ‘mere gifts’ we might call them.

1.1 Appraisal Respect

Kant’s stated views in the context of his ‘animal ethics’ are suspicious to say the least, and what is worrisome is that they are not peripheral to Kant’s positive moral doctrine but instead quite central to it, following directly and unequivocally from the fundamental principle of morality as described in the *Groundwork* and elsewhere. His famous ‘formula of humanity’, for instance, depicts morality as requiring that we treat *humanity* in persons as ‘an end in itself’ and “never merely as a means” (G 4:429), thus giving to understand the all beings lacking humanity (including the animals among them) are rightly treated as ‘mere means’.

I don’t want to jump to conclusions just yet about the credibility of Kant’s moral philosophy as stated, not before doing an effort to figure out what he might be trying to get at, or whether there is some important lesson to be learned in what he’s saying. I will now begin to discuss some of the other claims that Kant makes in relation to the moral standing

⁸⁵ As far as I know, Kant never explicitly argued for the moral permissibility of eating animals, but he did make a number of remarks seemingly entailing that he considered it perfectly permissible (e.g. MM 6:409). Kant is by no means alone in this, however; Jeremy Bentham (who some portray as an animal liberationist) argued for us being allowed to kill animals whom “we like to eat”, adding (Bentham 1823/1907, p.311): “we are the better for it, and they are never the worse. [...] The death they suffer in our hands commonly is, and always may be, a speedier, and by that means a less painful one, than that which would await them in the inevitable course of nature.”

of animals, hence trying to provide a larger picture and preparing the ground for the story that I came to tell. The following passage from the *Second Critique* (which I will be referring to as the ‘respect passage’) strikes me as the best place to start dialectically speaking.⁸⁶ Let me start by giving you the first line only (CPrR 5:76):

Respect is always directed only to persons, never to things. The latter can awaken in us inclination and even love if they are animals (e.g., horses, dogs, and so forth), or also fear, like the sea, a volcano, a beast of prey, but never respect.

Kant is here telling us that *respect* (important word) is owed only to persons (a.k.a. human beings), and so people take him to be denying the validity of the kinds of ideas of ‘animal dignity’ and ‘animal equality’ that utilitarians (and others who identify themselves differently but who nevertheless defend the moral standing of sentient beings in general) have been advancing. For if the proper object of ‘respect’ is that which alone has ‘dignity’ (G 4:435), and only persons (a.k.a. human beings) are proper objects of ‘respect’ (as the line just quoted suggests), then it follows that only human beings have ‘dignity’. Non-human animals are beyond the scope of ‘respect’, and this alone entails their lack of ‘dignity’.

But if non-human animals lack dignity and human beings (who are animals but not ‘mere animals’) have it, thus counting as the sole proper objects of respect, then it is not the case – what Utilitarians take to be the case – that all animals are equals in dignity and that therefore we ought to respect them all as equals: some animals don’t even have dignity, and those who have it (human beings) have it only on account of their humanity. If distinctively ‘animal dignity’ falls, as it does if humanity (the capacity for morality) is made the sole criterion for ‘dignity’ and ‘personhood’ and so on, the corollary notion of ‘animal

⁸⁶ We will see Kant advancing some claims and remarks concerning ‘respect’ and its proper objects, explicitly putting non-human animals out of scope of this attitude. However, I will be explaining why this passage in particular doesn’t in the least contribute (even by Kant’s own lights) to the exclusion of non-human animals from the scope of *respect* (in the sense in which people defending the notion of ‘respect for animals’ typically mean it). In other words: even though it may initially appear otherwise, and even though it is very often treated otherwise, the following passage is *not* part of the ‘negative’ story that Kant tells in his ‘animal ethics’ (which we certainly do well to clear up in order to better understand and assess that ‘negative’ story, thus the idea that this might be the best place to start dialectically speaking).

equality' also falls, and the related notion of 'respect for all animals as equals' is thereby also dismissed as inadequate, which of course is what Kant seems explicitly to saying in his famous 'respect passage'. But I ask: is that really the story that Kant means to be telling in this particular passage? I don't think it is, and Kant himself will tell you that it isn't. Let me give you the entire piece (the line already quoted included), and then I'll discuss it bringing back the relevant lines (CPrR 5:76-77):

Respect is always directed only to persons, never to things. The latter can awaken in us inclination and even love if they are animals (e.g., horses, dogs, and so forth), or also fear, like the sea, a volcano, a beast of prey, but never respect. Something that comes nearer to this feeling is admiration, and this as an affect, amazement, can be directed to things also, for example, lofty mountains, the magnitude, number, and distance of the heavenly bodies, the strength and swiftness of many animals, and so forth. But none of this is respect. A human being can also be an object of my love, fear, or admiration even to amazement and yet not be an object of respect. His jocular humor, his courage and strength, the power he has by his rank among others, could inspire me with feelings of this kind even though inner respect toward him is lacking. Fontenelle says, "I bow before an eminent man, but my spirit does not bow." I can add: before a humble common man in whom I perceive uprightness of character in a higher degree than I am aware of in myself my spirit bows, whether I want it or whether I do not and hold my head ever so high, that he may not overlook my superior position. Why is this? His example holds before me a law that strikes down my self-conceit when I compare it with my conduct, and I see observance of that law and hence its practicability proved before me in fact. Now, I may even be aware of a like degree of uprightness in myself, and yet the respect remains. For, since in human beings all good is defective, the law made intuitive by an example still strikes down my pride, the standard being furnished by the man I see before me whose impurity, such as it may be, is not so well known to me as is my own who therefore appears to me in a purer light. Respect is a tribute that we cannot refuse to pay to merit, whether we want to or not; we may indeed withhold it outwardly but we still cannot help feeling it inwardly.

Again, there is no doubt that Kant is here advancing claims about 'respect' and its proper objects, excluding non-human animals on the basis of the premise that only 'persons' (a.k.a. human beings) fall under the scope of this attitude. However, it is equally clear, I submit,

that Kant's claims in this passage concern a very particular *kind* of 'respect', and one that nobody – neither I nor anyone else I know of – is claiming for non-human animals (at least under the assumption that they are really *non-human*). We just have to realise that that the object of the kind of 'respect' that Kant is talking about in this context (known in the literature as 'appraisal respect')⁸⁷ is always someone – some human being – represented as *morally good*, or as having a high or positive moral stature: "before a humble common man in whom I perceive *uprightness of character* in a higher degree than I am aware of in myself my spirit bows" (CPrR 5:76–77).⁸⁸ But I am not making my case for animal dignity depend on the suggestion that our spirit should bow in moral admiration for non-human animals, obviously, and therefore I am not going to be bothered by Kant's 'respect passage', which according to some people I should be because – as they see it, presumably – Kant is here 'excluding animals from the scope of respect'.⁸⁹ No, not really. Nobody defending the dignity and respectability of animals (sentient beings) in general should be bothered by Kant's famous 'respect passage' because it doesn't even speak to the relevant effect: we can just read those lines and agree on the fundamental point.

But there are further problems that arise from portraying Kant's 'respect passage' in his *Second Critique* as adding to his negative story in the context of 'animal ethics'. Recall that in this particular context the proper object of 'respect' ('appraisal respect') is always someone regarded as *morally good*, which obviously puts *morally evil* human beings out of scope. Scoundrels (all of whom are *human beings*) are not suitable objects of this "tribute that we cannot refuse to pay to *merit*".⁹⁰ (CPrR 5:77). Kant concedes the point quite openly (CPrR 5:76): "A human being can also be an object of my love, fear, or admiration even to

⁸⁷ I am appropriating Stephen Darwall's terminological distinction between 'appraisal respect' and 'recognition respect' (Darwall, 1977).

⁸⁸ My emphasis.

⁸⁹ People who have appealed to Kant's 'respect passage' in this way include: Kain, 2010, p.210; Wood, 1998, p.197. However, it's not these direct appeals in written argumentation that worries me most, but the general lack of pertinent clarification in Kantian scholarship on the particular, and also how easy the passage is normally just added to Kant's 'negative' story in the context of 'animal ethics' without protest.

⁹⁰ My emphasis.

amazement *and yet not be an object of respect* [a.k.a. of moral admiration / ‘appraisal respect’]”.⁹¹

Hence: if non-human animals go out of the scope of ‘respect’ (unqualifiedly, as it were) on the basis of the truth of Kant’s claims in his ‘respect passage’, then so do those among us who fail to be morally good as well, which may just be *all of us*.⁹² But of course it is essential to Kantian ethics (the positive ethical doctrine) that *all human beings ought morally to be respected in their dignity as human beings*, that there is something like *respect for human beings as such and in general* and so quite irrespective of thoughts concerning the *moral worth* of their character and conduct: a kind of respect that has nothing essentially to do with applauding or bowing before morally good human beings. And this can only mean one thing: Kant’s ‘respect passage’ cannot be treated as giving us an *exhaustive* account of ‘respect’ and its proper objects – which is how it would have to be treated is if it is to be perceived as putting non-human animals unqualifiedly out of scope – *because that is not its purpose*.⁹³ As I have pointed out, Kant’s ‘respect passage’ is intended as an account of a

⁹¹ My own emphasis. The square brackets are also mine, of course, to make a contrast between ‘respect’ (which here I construe as ‘moral admiration’) and the kind of ‘admiration’ Kant is referring to in the line I just quoted in the main text, which (given the context of what he is saying) I construe as ‘wonder’ or ‘amazement’ with the non-moral perfections of a human being, e.g. her sheer skills as a footballer, something of the same general kind that may also apply in relation to non-human animals (e.g. amazement at the juggling skills of an utter).

⁹² If it would turn out that no real morality has ever been found in human action, as Kant himself entertains (G 4:407), then Kant’s ‘respect passage’ would entail that no human being has ever been a proper object of ‘respect’.

⁹³ Someone might insist on the relevance of ‘appraisal respect’ to the issues on the table by pointing out the following: even though evil human beings are not in fact worthy (proper objects) of ‘appraisal respect’, which makes it seem as if they are on a par with non-human animals in this regard, i.e. all of them *just failing* to be morally good, nevertheless they are *real candidates* for this worthiness: *all human beings* can in principle – at any point in their lives – become proper objects of ‘appraisal respect’ by making themselves morally good. In this, the argument goes, human beings are all quite different from non-human animals, and different *in kind* in fact, and different *in virtue of their humanity*, and different in a way that may strike us as *morally significant*. I don’t want to deny the connections here, nor do I mean to deny what is obviously true: a morally evil human being is to ‘appraisal respect’ what a ‘loser’ in a sporting competition is to the awarding of the winner’s prize, which is very different from how a non-participant stands in both scenarios: the one (the person on the

particular *type* of ‘respect’ (‘appraisal respect’) which exists and is present in our minds: that’s all there is to it. Kantians almost universally agree that there is *also* such a thing as ‘respect for *all* human beings’ as something that morality *requires* of us, and I am suggesting that we should include in the picture a similar (duty-involving) kind of *respect for animals* (*sentient beings*) *as such*.

There is an important thought that just suggested itself. The notion of ‘respect for animals’ that I am pursuing in this **Part II** is meant to pick out an attitude the instantiation of which (in action) is *morally required* of us, and it is a notorious feature of ‘appraisal respect’ (the kind of respect implicated in Kant’s ‘respect passage’) that it cannot be so required. ‘Appraisal respect’ is thought of as a *feeling* that we *naturally* or *inevitably* – if only as moral beings – have as a result of representing some person or deed as morally good, and it is plainly incoherent to think of something as a duty which one cannot avoid doing. “Respect is a tribute that *we cannot refuse* to pay to merit, *whether we want to or not*; we may indeed withhold it outwardly but we still *cannot help feeling it inwardly*” (CPrR 5:77).⁹⁴ But if the kind of ‘respect for animals’ that I am proposing is meant to feature as a moral requirement, and the kind of respect that Kant is talking about in his ‘respect passage’ (*viz.* ‘appraisal respect’) is not, and this by Kant’s own admission, then we have yet another powerful reason to oppose portraying Kant’s ‘respect passage’ as excluding non-human animals from the scope of the *relevant* kind of ‘respect’. The lesson here should be that the term ‘respect’ has more than one legitimate use in our moral discourse, and it is perfectly clear that ‘appraisal respect’ (the scope of which everyone seems to agree on) is *not* the relevant kind in the context of our discussion about the moral standing of non-human animals.

The kind of ‘respect for animals’ that I am trying to motivate is intended as a fundamental *moral requirement*, a form of what is known in the literature as *recognition respect*

stands) certainly fails to win the price, and the other (the animal walking around being completely oblivious to morality) certainly fails to be morally good, but that is because neither of them even bid for the relevant achievement. What I fail to see, however, is how this particular difference between human beings and other animals is going to undermine notions of ‘respect for animals’ of the kind that I and many other people are defending (which aim to protect their *beneficiary status* in the moral community and *in no way* their participation as moral actors, which we admit is out of place as soon as we call an animal *non-human*).

⁹⁴ The emphases are mine.

(Darwall, 1977). As I have it, this is the kind of respect that consists in *relating to a person (as a matter of duty) with full recognition and consideration for her existence as a person.*

1.2 Towards Recognition Respect for All Sentient Beings

A different story has been anticipated when it comes to ‘recognition respect’. Not only is there widespread agreement (among Kantians but also more broadly) that *all human beings* are worthy – i.e. proper objects – of this kind of respect (respect for their *humanity* quite irrespective of the actual moral worth of their conduct), thus making it clear that reference to ‘appraisal respect’ is not going to exhaust talk about ‘respect’ and its proper objects. It is also agreed that this kind of respect is *required by morality*, meaning that we are in the right kind of philosophical territory.

What is truly at stake in our disagreement (which we are supposed to be having in relation to the moral standing of non-human animals) is nothing but the *scope* of a particular type of respect (recognition respect) the exemplification of which constitutes – as Kantians have intended it in relation to human beings as such, and as I am intending it in relation to both human beings as such and sentient beings as such – a *moral duty* of ours. Again, it is only here that we get a genuine disagreement between Kant’s view and the one that I am proposing: while I take the fundamental ethical notion of ‘respect for persons’ (the one that features as a fundamental moral requirement, a.k.a. ‘recognition respect’) to *contain* that of ‘respect for animals (sentient beings) as such and in general’, Kant and many of his followers have made it their goal to defend a thoroughly anthropocentric philosophy equating the notion of ‘respect for *persons*’ with that of ‘respect for all *human beings as such*’, not to say ‘persons’ with ‘human beings’ and ‘dignity’ with ‘human worth’.

In saying that there is something like ‘recognition respect’ in relation to animals (sentient beings) as such, I don’t mean to suggest that we take the kind of ‘recognition respect’ that is enjoined by Kant’s ‘formula of humanity’ (respect for *humanity* in persons) and somehow extend the scope of it to cover all the other (non-human) animals. Again, what I am doing is positing the existence of *two distinct types* of ‘recognition respect’, one concerned with the existence of *human beings as such* and the other one concerned with the existence of *animals (sentient beings) as such* (including human beings *qua animals* or – what

is the same here – *qua capable of happiness*). Now, speaking of two *types* in this context may give to understand that there is a ‘duty of recognition respect for human beings’ over here and a ‘duty of recognition respect for sentient beings’ over there which may potentially (given that it is a ‘different type’ and so apparently a ‘separate duty’) come into conflict with each other, which I really don’t mean to suggest. The point is rather that there are two *aspects* to ‘recognition respect’ for persons as such: two phases to the personhood of individuals and two corresponding and clearly distinguishable ways of relating to them that are fundamentally required by morality (always to be realised *together*) under the one rubric of *respect for persons as such*.

In the final picture *all* ‘duties of respect for persons’ feature as *both* ‘duties of respect for sentient beings as such and in general’ and ‘duties of respect for human beings as such and in general’. That is the claim that I am making, and all the questions that this claim raises (which there are many) I will try to address as the argument goes. For now, the point is simply to state the claim as strongly as possible: not only is ‘respect for animals / sentient beings’ something real and irreducible, it is also something that permeates *the entirety* of our moral vocation, featuring as an essential aspect of *every* moral duty that we have. In other words, there is no moral duty that we have that is not a *duty of respect for animals (sentient beings) as such*. There is no moral duty that is not a duty of ‘respect for human beings as such’ either, but nevertheless the reference to ‘animality’ (here as individuals in their *capacity for happiness*) is always *equally essential*. We can preserve and try to make sense of Kant’s ‘formula of humanity’ *as it is*; we need only introduce a *new formula* – a ‘formula of animality’ formalising the principle of ‘respect for animals’ – and try to put the two together in the formulation of a *single principle* of ‘respect for persons’ / ‘respect for all persons in their dignity’. The success of this project will depend on us managing to introduce sufficient and (more importantly) *adequate* nuance into the picture: something like a *model* or *schema* putting all the fundamental pieces together into a single structure is to be expected as an upshot.

2. The Agapist Project in ‘Animal Ethics’

Let me remind you that Kant categorises non-human animals as ‘things’ as opposed to ‘persons’. This is closely connected with the rejection of recognition-respect towards these beings, of course, as the basic idea is that this kind of respect is (even as I define it) necessarily directed to *persons* (which for Kant only human beings count as). Nevertheless, it should be noted from the outset that Kant is not thereby doing away with all ideas of the special moral significance of non-human animals *vis-à-vis* inanimate entities, and so I will begin this section by describing Kant’s (as I would like to call it) *agapist* project in ‘animal ethics’: the project of making ‘love’ – as opposed to ‘respect’ – the critical attitude in our interactions with non-human animals. We see a hint of this in Kant’s ‘respect passage’, which we have discussed before (CPtR 5:76):

Respect is always directed only to persons, never to things. The latter can awaken in us inclination *and even love if they are animals* (e.g., horses, dogs, and so forth), or also fear, like the sea, a volcano, a beast of prey, but never respect.⁹⁵

Animals seem to be worthy – i.e. proper objects – of *love* in Kant’s view, then. Not so ‘respect’, to be sure, but nevertheless: it is not as if a cow and a car (or even a tree) belong in the exact same fine-grained moral category (let alone in the same ‘category of being’). For even though both belong in the general moral category of ‘things’ in terms of their basic moral standing (respectability / membership in the moral community / citizenship), nevertheless the cow is a proper object of ‘love’, whereas the car and the tree *obviously* are not, i.e. at least not of the kind that is relevant in this context, namely *benevolence* or *kindness*. So there you go: the reason Kant is prepared to dismiss the notion of ‘respect for animals’ and related notions, and to do so openly and without fear of failing to do justice to animals, is that he takes himself to be giving animals their actual due, namely ‘kindness’ or ‘love’.

Judging from experience, most Kantians seem to be perfectly happy with this kind of arrangement, and by way of defending Kant they can always point to his most ‘animal-

⁹⁵ The emphasis is mine.

friendly' moments (all of which emphasise the notion of 'love') as ones that captures all that we really want to say in the context of 'animal ethics' (LE 27:710):⁹⁶

Any action whereby we may torment animals, or let them suffer distress, or otherwise treat them without love, is demeaning to ourselves.

Even if 'respect' is out of the question in relation to non-human animals, Kant is bringing 'love' to the rescue and asking that we remain content with this (as he might see it) reasonable provision. I am not persuaded, however, but I have to be extremely careful explaining why I disagree. I'm not necessarily opposed to Kant's – as I am initially construing it – *agapist* project in 'animal ethics', and I will just admit that his most 'animal-friendly' line (the one quoted last) roughly captures the essence of what I myself want to say in the end: what morality requires of us in relation to animals is 'love', i.e. that we be *kind* or *benevolent* to them, that we relate to them on the basis of a fundamentally *caring* attitude. And it doesn't come as a surprise that I am (at least initially) approving of Kant's 'agapist' project in this way, since the ethical view I am developing and defending in this **Part II** (Utilitarianism) is – as I have been describing it – an *agapist* doctrine that most essentially enjoins 'practical love' (i.e. benevolence or kindness from principle) for all sentient beings.⁹⁷ In the end, however, I have to reject Kant's own story in this area.

Why is it that I am not on board, then? Simple. While Kant's 'agapist project' is offered as an *alternative* to the 'respect project' that he considers proper only in relation to *human beings as such*, my own agapist project in relation to animals (sentient beings) is one that is constitutive of the 'respect project' in relation to *persons in general*. I have made this point more than once, and it suffices on its own as a reason to reject Kant's project in 'animal ethics'. But let me first point to what looks like a more fundamental setback of Kant's approach in 'animal ethics'. Initially at least, it is fundamentally unclear what type Kant has in mind when he talks about 'love' as the critical attitude in relation to non-

⁹⁶ Not that Kantians need to point to anything in Kant to defend their philosophical point here, of course.

⁹⁷ I am interpreting Kant's project as 'agapist' because he points to 'love' (benevolence or kindness) as the critical attitude. But *agape* is in essence *love from duty*, the pursuit of someone's happiness as a *morally obligatory end*, and it is far from clear that Kant is giving animals this particular kind of love.

human animals. Even if we agree that Kant is using the word ‘love’ in this context to pick out something like ‘benevolence’ or ‘kindness’, i.e. an active concern for the happiness or well-being of animals, it still remains fundamentally unclear whether he is talking about the *right kind* of ‘love’; and if he is, whether his is drawing philosophical conclusions that are consistent with that. Is Kant talking about the kind that morality is typically thought to essentially involve – the one that is implicated in the *command* to love one’s neighbour? Or is he talking about some lesser kind (e.g. pathological love)?⁹⁸ If he is talking about ‘practical love’ or ‘agape’ (the kind that morality is thought essentially to involve), then he couldn’t possibly *in light of that* remain wedded to the rejection of ‘respect for animals’ and related notions, i.e. unless he is prepared to introduce more categories of moral standing in the picture. How can it possibly be the case that *one’s neighbour* (as one calls the object of ‘practical love’) is to be thought of as a ‘thing’ which exists as a ‘mere means to an end’? But then, of course, what is more likely is that Kant is giving non-human animals ‘pathological love’ (i.e. sensibility or sympathy-based love) only, reserving the ‘practical’ (reason-based) kind for human beings alone.

And so I ask: is that enough ‘love’ to make up for no ‘respect’? I don’t think so. What this does is once again excluding non-human animals from the scope of morality altogether, at most bringing them back – as Kant tries to do – through the idea that cultivating kindness for animals (in ‘pathological love’) is useful for strengthening ‘practical love’ (which is essential to morality but is directed to human beings alone). But of course we ought not to go around destroying all kinds of things if we wish to strengthen (and even directly to fulfil) ‘practical love’ in relation to our fellow humans, and this brings us back to the idea that animals are – for Kant – essentially on a par with ‘other things’ from the moral point of view. Again, we can love (care for) animals in a way in which we cannot love (say) cars, namely we can be *kind* or *benevolent* to them (in the literal sense of the term), but this seems to make no essential difference for Kant, which I find very hard to understand.

⁹⁸ Kant’s theory of love has it that there are two philosophically significant kinds in this context, two types of benevolence (active concern for someone’s happiness) that are possible for human beings given their dualistic psychology. ‘Pathological love’ is love based on *sensibility*, and in particular in the *sympathy* that human beings ‘naturally’ have for subjects in their happiness and unhappiness; ‘Practical love’ is love (benevolence) *from duty* – the pursuit of someone’s happiness as a *morally obligatory end*.

He is not saying (what one would imagine he would say having given animals ‘love’ as *opposed to* ‘respect’, or nothing in particular): treat human beings as ‘ends-in-themselves’ (respect + love), the other animals as ‘neighbours’ or – to go ‘less radical’ – ‘darlings’ (love), and the rest of nature as a mere resource. He is saying: treat human beings as ‘ends-in-themselves’ and the rest of entities as ‘mere resources or supplies’, but do please treat all of the latter gently (*especially the animals*, because this comes much closer to our hearts) in order to fully treat human beings as ‘ends-in-themselves’, who alone deserve to be treated on the basis of truly moral attitudes like respect *and* practical love, who alone count as potential victims of theft and murder, and so on and so forth. Kant recognises that animals are special in that they, and not ‘other things’, are proper objects of ‘love’, but he makes no important effort on this basis to give animals – i.e. non-human animals – any truly special place in the moral universe. He doesn’t give them any kind of *citizenship* in the moral community on this basis, certainly.

Let me go into more detail to have clarity on the options available. We can appeal to Kant’s distinction between two kinds of love / benevolence to distinguish between two possible views that one can endorse (and I don’t mean as an interpretation of Kant, but simply *for oneself*): one saying that in relation to non-human animals only ‘pathological love’ is possible (then somehow – but only indirectly – making it part of morality), and another also affirming ‘practical love’ (i.e. love from duty, love grounded in fundamental moral principles, love that is itself essential to morality). We can bring it down to the antagonism between the ‘practical-love view’ and the ‘no practical-love view’, which, in premutation with the antagonism between the ‘respect view’ and the ‘no respect view’, between the view that affirms ‘respect for animals as such’ and the one that denies this notion, generates four possible views that one can endorse:⁹⁹

⁹⁹ My position is (2) and the one most likely to be Kant’s is (3), both with their heading in bold. Of the other two, only (4) is of significant interest for my argument.

[~PL + R] One is the view that denies ‘practical love’ for animals but nevertheless affirms the notion of ‘respect for animals as such’. This view cannot be attributed to Kant, nor is it the view that I am defending.¹⁰⁰

[PL + R] I will devote all my attention to the remaining three views, only one of which (the one that I am proposing) affirms the notion of *respect for animals (sentient beings) as such*. The view also affirms the notion *practical love for all animals*, and takes up the task of explaining how the two notions are related. For obvious exegetical reasons, this view (affirming both positives) cannot plausibly be attributed to Kant. I could of course insist that the view deserves – as I myself intend it – the label ‘Kantian’, if only side by side those of ‘Agapist’ and – as I also intend it – ‘Utilitarian’, but that is of course the argument that I am supposed to be giving in this work as a whole.

[~PL + ~R] The third view is the one that denies both ‘practical love’ and ‘respect’ in relation to animals as such and therefore also in relation to non-human animals. It gives the latter ‘pathological love’ to be sure, but that and nothing more. This is the view that is most plausibly attributed to Kant, I think, since it coheres better than the others with his insistence on calling non-human animals ‘things’ and the other disdainful statements that he makes in relation to them: e.g. that we can go ahead and take their skin (and perhaps even their flesh, and so by implication their lives) with a sense of having a ‘right’ to do so, with a sense of taking something that belongs to us as a commodity (statements that I nevertheless take to be in some tension – at least in spirit – with some of the things that Kant says in relation to animals, including his remarks giving them – if only always pathological – love).

[PL + ~R] It sounds almost incoherent to declare animals ‘things’ with no worth beyond ‘price’ (usefulness) – which is how one seems forced to talk about them if one denies them ‘dignity’ and ‘respect’ – and yet think of them as falling under the scope of ‘practical love’:

¹⁰⁰ The reason this option exists is simply that it was generated by the permutation that I suggested to produce the other three views on the table, which alone I will be discussing.

to first portray them as *commodities* and then include them among *one's neighbours* or *co-beneficiaries of morality*. Nevertheless, that is what appears to be the fourth possibility on the table, which we may want to attribute to Kant (for the sake of having a more 'animal-friendly' reading of his moral philosophy) even in the face of his denigration of animals.¹⁰¹ This is the view that many contemporary Kantians have adopted (e.g. Denis, 2000; Timmermann, 2005) in an attempt to prevent portraying concern for animals as depending for its justification on contingent facts about human psychology and society. To be clear, I don't think these people are actually committed to calling animals 'things' (because they acknowledge there is something very odd about this) but are instead committed to them having a special kind of moral standing - which as far as I know they haven't named - putting them *above 'things' and below 'persons'* (or - not to suggest that the point is to convey a simplistic hierarchy - at least distinguishing them from 'things' but not quite making them 'persons').

At this point in the argument, the main goals for me have become: (i) to elucidate the nature of a certain *duty of love (practical love) to animals* that I think we have, and (ii) to show that this *duty of love* entails 'recognition respect' for animals (sentient beings) as such. I also have to (however briefly) explain *why* we have the duty in question in the first place, of course. Putting everything together, the goal is to argue for the existence of a respect-entailing duty of 'practical love' for *all animals* (i.e. respect-entailing in relation to the animals themselves).

What I will do, then, is offer a critique of the so-called 'indirect view' that Kantians have tended to favour in their proposed accounts of our duties in relation to non-human animals: the view that our duties with regards to animals are not 'duties to animals' (in the same way that our duties with regards to cars are not 'duties to cars'). This target view, I

¹⁰¹ The thought that this view may plausibly be attributed to Kant (even in the face of his denigration of animals) is motivated more by philosophical than merely exegetical considerations. Maybe we have to agree that Kant himself never gave his assent to the notion of 'practical love' in relation to non-human animals, but he got so close, or his views allowed for it so perfectly (to the point of calling for it when one looks at the big picture), that one may just take the next step on his behalf. Similarly with the notion of 'respect for animals', which I argue is entailed by the notion of 'practical love'.

take it, is indistinguishable from the ‘no respect view’, only more accessible to criticism (in virtue of the unpacking of concepts that it does, telling us what ‘recognition respect’ actually amounts to). If I show that we have duties to animals (duties to individuals *as capable of happiness*, duties in which sentient beings *as such* figure as *moral patients* or *direct moral considerables*), I will have done all that can possibly be asked by way of justifying the notion of ‘respect for animals /sentient beings as such’, and those of ‘animal dignity’ and ‘animal equality’ (the dignity and consequent equality of all sentient beings as intended beneficiaries of morality) will follow. The notion of ‘practical love’ remains central to the argument, of course, and even more so now as we begin to move into the territory of stating *Utilitarianism* as an ethical theory.

3. The Notion of ‘Duties To Animals’

The only proper argument that Kant gives against the notion of ‘duties to animals’ – on the ground that only *duties to human beings* are possible – appears in a section of the *Doctrine of Virtue* entitled ‘On an **amphiboly in moral concepts of reflection**’. I will give you the entire piece (conveniently edited to make it shorter and also to separate the two parts that it has) so you can get a good sense of what Kant has in mind in this context; then I will discuss it in detail quoting the relevant lines again (MM 6:442–443):

As far as reason alone can judge, a human being has duties only to human beings (himself and others), since his duty to any subject is moral constraint by that subject’s will. Hence the constraining (binding) subject must, *first*, be a person; and this person must, *secondly*, be given as an object of experience, since the human being is to strive for the end of this person’s will and this can happen only in a relation to each other of two beings that exist (for a mere thought-entity cannot be the *cause* of any result in terms of ends). But from all our experience we know of no being other than a human being that would be capable of obligation (active or passive). A human being can therefore have no duty to any beings other than human beings; and if he thinks he has such duties, it is because of an *amphiboly* in his *concepts of reflection*, and his supposed duty to other beings is only a duty to himself. He is led to this misunderstanding by mistaking his duty *with regard to* other beings for a duty *to* those beings.

[...]

With regard to the animate but nonrational part of creation, violent and cruel treatment of animals is far more intimately opposed to a human being's duty to himself, and he has a duty to refrain from this; for it dulls his shared feeling of their suffering and so weakens and gradually uproots a natural predisposition that is very serviceable to morality in one's relations with other people. The human being is authorized to kill animals quickly (without pain) and to put them to work that does not strain them beyond their capacities (such work as he himself must submit to). But agonizing physical experiments for the sake of mere speculation, when the end could also be achieved without these, are to be abhorred. – Even gratitude for the long service of an old horse or dog (just as if they were members of the household) belongs *indirectly* to a human being's duty *with regard* to these animals; considered as a *direct* duty, however, it is always only a duty of the human being to himself.

Something that should be perfectly clear – from just reading the text – is that Kant is not denying the existence of duties to *treat animals well*. After all, the whole philosophical point that he is making in the context of his argument concerns the *nature* of those very duties: what he is saying is that our duties to treat animals well, which exist, are *not* to be thought of as 'duties to animals' but only as 'duties to *human beings* with regards to animals' – duties to humanity that of course relate to animals in that they are the duties that govern our *interactions with animals*. And not only that. Kant is even giving us a list of some such duties that he thinks we have. He says of a certain attitude that he thinks we ought to have (gratitude for the services of horses and dogs) that it “belongs *indirectly* to a human being's duty *with regard* to these animals” (MM 6:443). He also says of a duty that he thinks we have (the duty not to perform experiments on animals for purposes of sheer speculation) that failure to honour this duty is “to be abhorred” (MM 6:443).¹⁰² And the list is completed by a couple of constraints on the way we ‘use’ animals for our ends and benefit: that we are allowed (“authorised”) to kill animals (for purposes like eating them, presumably) but only

¹⁰² I don't fully appreciate the claim here (let animals *never* be taken as experimentation material), but I do very much appreciate the language.

quickly and without pain, and that we can put them to work but only in ways that don't overload them (MM 6:443).

That is as pertains to the second part of Kant's passage as quoted, where he is doing an effort to explain the *practical implications* of the view he is proposing. The argument that Kant gives for the view itself (his so-called 'indirect view') is condensed in the first part of the passage, which I will be quoting again in a second. Before doing that, I need to go over two preliminary points concerning two (crucially important) dialectical facts that emerge in relation to second part of the passage:

(1) Kant's proposed list of duties with regards to animals (in the second part of the passage above and elsewhere in his work) is bound to strike many people as falling short of our true demands in this context. However, this is not the place to correct Kant on the particular (at least not in any significant detail, and certainly not in the mood of arguing against people who would declare themselves sceptical to begin with). Here we are not as concerned with 'the list' as we are with the nature (or the structure) of the duties that make up the list. After all, it seems possible – dialectically speaking – for people who see the need of a more robust list of duties to treat animals well to remain faithful to Kant's philosophical analysis (his so-called 'indirect view'). They can resist revising Kant's more foundational claims in philosophy while providing more satisfactory 'content'. In light of this dialectical fact, it is proper for me to bracket (i.e. for the most part, and at least for the time being) the issue of 'content' (meaning 'the list of things that we ought to do' without looking very much at their nature in the more philosophical sense) and focus on the philosophical analysis *itself*. Just to be clear, I will be assuming that *veganism* (here as the keeping of all animals outside the palate of resources or supplies at our disposal) is one of our general duties in this context.

(2) Perhaps the best way to state the basic disagreement between Kant's view and the one I am proposing is to notice what is at stake in terms of conceptualising our moral relations with non-human animals. If your preferred strategy in developing 'Kantian animal ethics' (i.e. if that is what you are doing) is to endorse the 'indirect view' and revise it only at the level of practical prescriptions, i.e. in terms of the list of duties that we have *in their extension*, you can certainly think of violations of those prescriptions as rendering our actions *immoral*; what you won't be able to say, of course, is that animals are thereby *wronged*

in their dignity as persons. If you think that ‘veganism’ is among our duties, for instance, but agree with Kant on its not being a ‘duty to animals’, then you won’t be able to say that violations of it are such that animals are thereby *wronged* or *disrespected* (at least not in the relevant sense of the term). You won’t, for instance, be able to say that killing animals unlawfully (morally unlawfully) counts as *murder*. But I do want to affirm – what I admit is a painful truth – that in our practice of killing animals (for food, for instance) we are committing the sacrilege of murdering our fellow persons. Accordingly, I need to revise Kantian ethics in a much more fundamental way (i.e. provided that I want to claim the label ‘Kantian’ for myself) if only to be able to call things by their name; to call things ‘murder’, and ‘rape’, and ‘genocide’, when I think they are.

Knowing what is at stake (generally speaking) in our ‘dispute’, we can finally move to the analysis of Kant’s more technical statement of – and argument for – the so-called ‘indirect view’. Let me quote the relevant piece again (MM 6:442):

As far as reason alone can judge, a human being has duties only to human beings (himself and others), since his duty to any subject is moral constraint by that subject's will. Hence the constraining (binding) subject must, *first*, be a person; and this person must, *secondly*, be given as an object of experience, since the human being is to strive for the end of this person's will and this can happen only in a relation to each other of two beings that exist (for a mere thought-entity cannot be the *cause* of any result in terms of ends). But from all our experience we know of no being other than a human being that would be capable of obligation (active or passive). A human being can therefore have no duty to any beings other than human beings; and if he thinks he has such duties, it is because of an *amphiboly* in his *concepts of reflection*, and his supposed duty to other beings is only a duty to himself. He is led to this misunderstanding by mistaking his duty *with regard to* other beings for a duty *to* those beings.

The general thought that Kant wants to convey is that the notion of ‘duties to animals’ is a kind of *illusion* or *deceptive appearance* (an *amphiboly* as he calls it). We can only have duties to persons (a.k.a. human beings) who are objects of experience (whose presence we encounter), since only human beings are capable of active obligation (i.e. of obligating) and passive obligation (i.e. of being obligated). Kant recognises that at least some people take themselves to have ‘duties to animals’ (thus the purported amphiboly), but he tells them

that if they were to get in a more philosophical mood – and think about the actual grounds of their duties and the structure of their own moral relationality – they would understand that it is all about honouring their own humanity (self-legislation) and that of other human beings (co-legislation), not at all about honouring the animals they are interacting with (whom of course they are obliged to ‘treat well’). We can break down Kant’s argument – the relevant features of it – for closer examination as follows:

- (1) Having a duty to some subject means *being constrained (obligated) by that subject’s will*.
- (2) Animals (non-human animals) lack a will of the kind (a morally rational will) that is capable of imposing moral constraints, i.e. they lack a will capable of obligating.
- (3) Therefore, we cannot have *duties to animals*.

People who want to take this argument seriously but want to reject its conclusion – because like me they hold belief in the existence of ‘duties to animals’ – can appeal to a number of strategies. They can reject premise (1) and try to show that Kant is wrong in thinking that ‘having duties to someone’ means, or can only mean, ‘being morally constrained by that someone’s will’ – that ‘being obligated to’ – in a way that we recognise as essentially belonging to morality – doesn’t necessarily mean ‘being obligated by’. People who want to take this strategy can proceed by explaining how it is that we can have duties to animals (i.e. to sentient beings as such) without appeal to the idea that their will has the power to impose moral constraints on ours, or that we are obligated by them. People can also question premise (2) in the argument, which they can do by trying to show that despite lacking moral rationality ‘animal wills’ are still capable of imposing moral constraints on us, in some *distinctive* way that has to be specified. I guess someone could reject both (1) and (2), although I would expect most people to reject only one premise in this argument. As you might have guessed, my own strategy revolves around rejecting premise (1). The reason I think Kant’s argument fails is because I reject the claim that ‘having duties to someone’ means, or can only mean, ‘being constrained that that someone’s will’.

Now, it might initially seem that rejecting premise (1) in Kant’s argument commits me to denying that ‘having someone’s will impose moral constraints on one’ amounts to ‘having a duty to that person’. If you think about it, however, it doesn’t. Rejecting the truth

of Kant's first premise – as I am doing – is *not* the same as denying the *positive connection* that it posits, between 'having a duty to someone' and 'having that someone's will impose moral constraints on one'. In fact, I am agreeing with Kant that such is one of the valid meanings of the locution 'having a duty to someone': you can, for instance, think of morality (taken as a whole) as something that you owe *to yourself* because (or perhaps it would be better to say: *in the sense that*) you are morally obligated to *by your own will*. What I am rejecting, rather, is the far more ambitious claim that such is the *only* meaning that it has – or (to be more precise on how exactly I read the claim) the only valid meaning that it has *and that we recognise as essentially belonging to morality*.

I haven't yet told you what the 'second meaning' is of the locution 'having duties to someone' (directed duties) which figures in our notion of 'duties to animals', but first I need to tell you what I don't mean when I claim that such duties exist. In arguing that there is such a thing as 'duties to animals', I don't mean to be locating the source of moral normativity or the ultimate ground of moral duty in sentient life as such, neither in happiness (its distinctive 'good') nor in self-love (its internal practical principle). Likewise, in arguing for Utilitarianism (the moral doctrine that enjoins impartial concern for all sentient creatures), the idea is not to make the morality of actions depend on their felicitic consequences, neither partly nor indirectly, neither in actuality nor in expectation. As I am portraying them, remember, happiness and utility (conduciveness to happiness) fail to be source or ground of *any* moral normativity. All by itself, happiness is a *moral zero* (which is not to say that 'it doesn't matter' because surely it matters in many ways, but simply that *all by itself* it fails to generate any kind of *moral incumbency* upon the human will).

If the happiness of subjects is morally incumbent upon us (which undoubtedly it is), it is only *because the moral law (human reason) speaks in favour* of – in the sense of *demanding* – relating to those subjects in a certain way. Now, following my initial assumptions in this work, I take it as a *given* (a fact that is *known to begin with*) that the happiness of animals in general is an unconditionally necessary object of the human will, something *towards which* our will (our intentional activity) ought unconditionally to be (outwardly) directed, but the explanation for this normative fact is not to be found in happiness itself, but in the moral prescriptions of our own intelligence. It's our own existence as morally rational beings that gives happiness (everyone's happiness) the kind of *moral incumbency* that we all assume it

has, which we have in mind when we think of it as something we ‘*ought morally* to do something about’. I am identifying this feature of happiness (which other things can also have, although perhaps not exactly in the same way) as the *moral incumbency* that we acknowledge it to have, although we can also talk about the *obligatoriness* that it has as an *end* for us. Or – if you will – we could also follow the somewhat extravagant ways of modern philosophy – which I find surprisingly useful – and describe this normative property of happiness (of everyone’s happiness) as its ‘ought-morally-to-do-something-aboutness’.

By the way, I don’t think one needs to give an elaborate argument for this Kantian point about the source or ground of the moral incumbency of happiness as an end for us (although you can of course avail yourself of the argument given in **Part I** against what I called ‘the standard utilitarian formula’). One need only ask: how can happiness (or anything else for that matter) have ‘*ought-morally-to-be-pursuedness*’, say, but for the existence of a *moral ought* addressing a suitably responsive will and thereby putting this will in the relevant normative relation with the happiness in question? Obviously, it cannot happen. And if we think, as we must (because it is fairly obvious), that moral oughts in general cannot exist in the absence of moral rational beings, then we are committed to thinking that the moral incumbency of things upon the activity of morally rational beings (upon their will) is grounded in the very existence of such beings, i.e. on the very existence of their will, which is then assumed to have the power of imposing moral duties upon itself. Moral oughts are generated by our own reason in its capacity to demand that certain things be done and certain others be left undone, and thus the sole origin of the moral burden that we bear as morally rational beings. If we care morally about anything, therefore, we do so only because we are obligated by the demands of our own reason.

But wait a minute. Saying – as I am doing – that we care morally about the happiness of animals ‘only because’ we care morally about living up to the demands our own humanity seems very much like toying with the ‘indirect view’ which Kantians have tended to favour, a view which I mean to be attacking instead of defending. I am saying that the *moral incumbency* of things in general is grounded in the prescriptions of our own reason, and this claim doesn’t discriminate between animals and other kinds of entities, which is exactly what a Utilitarian mind is supposed to find troubling: the ‘*ought-morally-to-be-protectedness*’ that attaches to animals is just as uniquely grounded in human reason as

the ‘*ought-morally-to-be-protectedness*’ attaching to, say, churches (the buildings). Well, this is of course a ‘traditional Kantian’ claim that I am making, I admit, but you would be utterly confused if you took it as a statement of the so-called ‘indirect view’. I agree with Kant’s claim that human reason is the sole *ultimate ground* of moral normativity, that the *moral incumbency* of things in general upon our will is grounded in the moral prescriptions of our own reason. I just don’t see how the telling of this emancipating story about the source of moral obligation could – in a philosophically respectable way – turn into a story of *exclusion* – which of course is how I see it – of non-human animals from the domain of respectability or direct moral considerability; I mean, beyond simply saying that we ‘owe nothing to them’ by way of obeying or recognising their moral authority, which of course they have none to be respected or recognised.

Let me continue in the mode of reflection. If my own reason is that which alone puts me under obligation and makes things in general morally incumbent upon my will (that which alone morally obligates me), does that mean I owe nothing to animals by way of moral action? That if I were to mistreat them I wouldn’t be betraying them? I just don’t see how it could. I have seen many claims about how human beings are distinct from other animals, how human beings are special in ways that are of genuine moral significance: that they alone are morally rational, for one, which entails that they alone are bearers of moral duties, and that they alone are to be treated as themselves having moral duties, and in fact – to talk about the specific point we are discussing right now – that they alone are capable of *morally obligating* (of generating moral obligations by imposing them upon themselves). The truth of these claims I can see very clearly; what I don’t see is how this truth could possibly undermine the notion of *duties to animals*. Again, we can agree that the existence of human beings is alone responsible for introducing a moral dimension into the world, but this is a very technical statement (albeit a very elemental one) that merely opens the door for all substantive moral theorising instead of bringing it close to an end: it is certainly not going to prevent us from introducing morally charged language to describe the moral and immoral qualities of our interactions with non-human animals, nor is it going to stop us from thinking that non-human animals can in all kinds of ways be *wronged* and *betrayed*.

Let me try to give a kind of argument on the basis of what has been said so far. Granted that the moral incumbency of both churches (the buildings) and non-human

animals upon our will (as entities that we *ought morally* to protect, for instance) is equally grounded in the prescriptions of our own reason, it can still be the case that we have to ascribe to non-human animals, and not so to churches, ‘ought-morally-to-be-respectedness’ and ‘ought-morally-to-be-protected-for-their-own-sakeness’. In other words, it can still be the case that we have duties *to animals* and not so to churches or other inanimate entities. We can understand the point made by this argument as follows. Moral oughts can be connected to all kinds of things, and no matter what connection we have in mind it is always going to be the case that moral oughts are grounded in the prescriptions of our own reason; but the specific attitudes (whose instantiation in action) morality requires of us by means of her moral oughts are *directed to* very specific kinds of things, i.e. they have very specific *direct objects*, which means that we *do* have to discriminate between different kinds of objects (some of which are of course *subjects*) when it comes to thinking about the *directedness of our moral duties*, this being the specific kind of connection that is at stake in our discussion. Practical love, for instance, is exclusively directed to *sentient beings* as intended beneficiaries, meaning of course that all duties of love are *duties to (duties towards) sentient beings as intended beneficiaries*.

The ‘indirect’ / ‘no duties to animals’ / ‘no respect’ view which the majority of Kantians favour in the context of ‘animal ethics’ is intended quite explicitly to prevent a kind of *false reductionism* that Utilitarians have seen themselves tempted to fall into: the idea that our mistreatment of animals is wrong or (more to the point) *immoral*, not in virtue of something going on within ourselves as moral beings (something having to do with the intrinsic quality of our own morally-informed volition), but in virtue of facts about the well-being of animals. Altman, for instance, has protested that the claim (which Utilitarians and others often make) that animals should be considered *directly* “is meant to make sure that animals are considered even when our moral characters are not affected.” (Altman 2019, p.280). I don’t like to keep rehearsing points that I have made more than once, but I really feel like I have to: *of course* we have to accept that if some action in relation to animals fails to render *us* (i.e. our own will) moral or immoral, morally good or morally evil (valences that I take to be irreducible to the felicific in any form), then it just follows that such an action falls outside the scope of moral assessment altogether. The question in the context of ‘animal ethics’ is not, say, whether harming animals really harms animals: of course it

does. Nor is it whether, say, harm done to animals counts as an intrinsically bad thing: of course it does, at least in some sense – surely, for one, in the strict *felicific sense*, which of itself entails nothing morally practical (no ‘moral-oughtness’) but is not for that reason ‘unimportant’. The real question is whether, say, there is something to harming animals, something *about us* connected with the harming of animals, which could perhaps *render us and our conduct (as human beings) immoral*.

Altman makes it seem as if people endorsing the ‘direct duty view’ (the contender of his own ‘indirect view’) are not even addressing the right question. But I just raised the right question, and I reject the ‘indirect view’ in favour of the ‘direct’ or ‘duties to animals’ view. He also makes it seem as if people endorsing the ‘direct duty view’ are committed to explaining the immorality of our mistreatment of animals in terms of something happening to animals themselves, such as the harm (the pain and the privation of happiness) we inflict on them. But I am not committed to any such claim. In fact, we needn’t even be talking about *harm* in this context, since it has already been established that immorality always features as a failure of *attitude* on the part of a morally rational being (see **Part I**), a failure that may or may not translate into harm. I have even pointed to the relevant (relationally-specific) attitude in this context: it is the failure of *practical love* in our interactions with animals that renders our actions immoral, which it does *directly* and therefore quite independently of all felicific properties that may attach to our personality and conduct. If this is the critique coming from the ‘indirect view’, I cannot but think that people advancing this critique – and endorsing the view – are missing the point. These Kantians may be right in pointing to a peculiar kind of false reductionism that Utilitarians have been tempted to fall into (although it would have to be shown who among them – and how and when exactly – they actually fall into it), but I don’t think they are saying anything that could pose a threat to the notion of *duties to animals* (i.e. to non-human animals) and related notions. This will take me some effort, but I would like to take the offensive and argue that

that ‘indirect view’ in ‘animal ethics’ can be shown to be guilty of a *false reductionism* of its own, and that therefore it is false.¹⁰³

Saying – as I am doing – that bad felicitic consequences are morally immaterial in that they don’t contribute, let alone ground, the *immorality* of our mistreatment of animals, is not the same as, nor does it entail, denying the existence of *wronging relations* involving animals as victims. Likewise, saying that the benefits (the happiness) that animals derive from our treatment of them doesn’t as such contribute to, let alone ground, the moral goodness of our conduct, is not the same as, nor does it entail, denying that morality essentially involves *acting for the sake of animals*. In order to show that I am wrong here and that one thing really follows from the other, that non-human animals failing to be the ground of anything moral entails their non-participation (not even as objects or recipients) in fundamental moral relations, it would have to be shown that reason (the admitted ground of our moral duties) fails to generate a ‘duty to animals’, that it (reason) fails to put us in special kind of moral relation to sentient beings as such which we recognise as essentially belonging to morality. I will be giving a chance to the ‘indirect view’ to try to do exactly what I am saying it has to do, but not before beginning to develop a more elaborate case for my own view.

4. ‘Kant’s Dictum’, Practical Love for Animals, and Self-Givingness

There is a crucial question that I haven’t yet addressed: Why is it (normatively speaking) that we have to care for the happiness of animals in the first place? Independently of what we think about the nature of our duty to ‘treat animals well’, to be practically concerned with their happiness, why do we even have this duty to begin with? Why should we even care? Believe it or not, the answer that I want to give to this question is an adaptation of one of Kant’s dictums: “I want everyone else to be benevolent toward me (*benevolentiam*); hence I ought also to be benevolent toward everyone else.” (MM 6:451). The point can be

¹⁰³ I will be substantiating this charge of false reductionism, but I have to warn you that it will take me some time to prepare the ground for that argument. The argument is closed in the only (long) footnote in subsection 5.5 below.

made in a more elaborate way by saying – as Kant also says – that “reason [...] *permits* you to be benevolent to yourself on the condition of your being benevolent to every other as well.” (MM 6:451) I will focus on discussing the first of these lines, which I will be referring to as ‘Kant’s dictum’.

4.1 Kant’s Dictum

It is my view that, unless we put an emphasis in the wrong place with the explicit purpose of excluding non-human animals, a proper analysis of Kant’s dictum (‘I want the benevolence of others, therefore I ought to be benevolent to every other’) is bound to lead us to the conclusion that *all animals* fall under the scope of our duty generalised benevolence or practical love. Kant himself presumably disagrees, remember [~PL + ~R]. In relation to non-human animals, Kant says: I don’t owe benevolence – in the sense of *practical love* – to these animals, but only to human beings, because I cannot even want them to be benevolent to me, since they can only (in virtue of their own ‘brute constitution’) be benevolent to themselves.¹⁰⁴ If I cannot even have an interest (the initial interest that is registered in Kant’s dictum) in the benevolence of animals because such is *impossible*, then (the argument runs) there is no contradiction in me denying animals my benevolence: there is no imbalance or free-riding, no ‘exceptions for oneself’, no violations of the requirement of ‘reciprocity’, in denying ‘giving’ when one cannot even entertain the

¹⁰⁴ The difficulty arises from the very start that we may not have clarity about the exact meaning of the claim (which I am putting Kant to say in this context on the basis that he excludes non-human animals from the scope of practical love) that *we don’t owe benevolence to animals*. Is it simply that the benevolence we ought to have to them is not *owed to them*? Or is it that we need not even be benevolent to them to begin with? Or is it that we ought to be benevolent to them, but since we ‘don’t owe it to them’ we can take this benevolence *less seriously* than we take benevolence to human beings (to whom the benevolence is owed in reciprocity)? These questions relate to the question what kind of ‘love’ it is that Kant gives to animals, which I have concluded is *only* pathological or sympathy-based. Now, it is always hard to pinpoint exactly what it is that Kant denies to animals at any particular point, but I will be assuming that he goes as far as denying benevolence itself (at least in the form of ‘practical love’) to non-human animals.

possibility of ‘receiving’ from the relevant party. Or so, again, the thought goes (presumably, for the moment).¹⁰⁵

I want the *benevolence of others*, therefore I ought to be benevolent to every other. Who is the ‘every other’ here? Well, if you put the emphasis on the *benevolence of others* (as I just did just to illustrate what Kant *may* have in mind which leads him to exclude non-human animals from the picture) could lead you to think that it is right and proper to limit the *scope of benevolence* (one’s own requirement of practical love) to the domain of *human beings*, who alone can be universally benevolent (or at the very least benevolent to individuals *other than themselves*) and whose benevolence alone I can want while holding it to be possible: more importantly still, whose benevolence alone I can consider in relations of *reciprocity* with mine. To the human being whom I am helping, I can say: I know you would do the same for me! To the non-human animal, however, I just cannot say that. So, if non-human animals don’t even figure in the *matter* that I plug into the formal principle of reason (i.e. my initial desire for everyone else’s benevolence, which given a basic recognition of the nature of things can only be a *desire for the benevolence of those who alone can be benevolent to me, and with whom alone I can therefore enter into relations of reciprocity*, a.k.a. human beings), how can I even entertain the notion of benevolence to these animals as a fundamental moral requirement incumbent upon me? How indeed. Can I justify to myself being indifferent (in the sense of ‘practical love’) to the happiness of a pig, say, because ‘given his brute constitution’ he cannot be benevolent to me? Here I just have to say (and I take responsibility for this) that I find the thought completely absurd. I cannot justify to myself being indifferent to the happiness of human babies on the grounds that they cannot be benevolent to me, and *just like that* I cannot justify indifference to the happiness of non-human animals on the grounds that they cannot be benevolent to me. Can I even justify ‘greater benevolence’ (benevolence greater in kind or even in degree) for a human being

¹⁰⁵ I will continue invoking ‘reciprocity’ in subsequent paragraphs in the context of my argument for all animals falling squarely under the scope of ‘practical love’ (our duty of ‘generalised benevolence’), but I will come back to this theme (now under the heading of ‘Contractualism’) in the context of my argument for the notion of ‘respect for animals’ (which constitutes my final argument for the notion of ‘duties to animals’, and in fact for the notion that ‘we owe’ moral activity *as a whole* to animals). I do more of an effort to do proper justice to ‘Kantian reciprocity’ in that context than I do in the present subsection.

vis-à-vis a pig because only the former can be benevolent to me or because only the former can stand to me in relations of ‘reciprocity’? Again, I don’t see how I could: practical love (benevolence from principle) is the kind of attitude that is blind to the moral credentials of its object, assuming only that it (i.e. the object, which of course is always a *subject*) be capable of happiness.

If there is something truly uncontroversial in our discussion (in terms of the meaning of basic concepts) is that the notion of ‘benevolence’ / ‘kindness’ equally covers *all animals* under its scope. We are agreeing that ‘benevolence’ is defined as *practical concern for someone’s happiness*, that is, and ‘animals’ (sentient beings) are exactly those beings who are *capable of happiness*. Therefore, all animals fall under the scope of benevolence. I don’t take any Kantian to be denying this fact. Kant certainly doesn’t deny it, since he allows for the notion of ‘pathological love’ (i.e. sympathy-grounded *benevolence*) in relation to non-human animals. In principle, therefore, without yet thinking of benevolence (kindness) as a moral requirement, let alone as a fundamental one, it is perfectly clear that we can be benevolent to a pig and a dog and a cat and a human being, but not so to a table or a car or painting or a tree. We can only be benevolent to *sentient beings*. But if this so, i.e. if benevolence – considered just as benevolence – is properly directed to non-human animals because in fact it is always essentially directed to *animals (sentient beings) as such*, how is it that when we make it into a moral requirement (when we think of it as directly incumbent – as a duty – upon our will) all of a sudden it isn’t?

Well, again, we can certainly point to the incapacity of non-human animals to be benevolent to us, which might get in the way because of some concern about ‘reciprocity’: perhaps that it would be *unfair* to care for someone as an equal who doesn’t have duties himself (who would count as something like a ‘free-rider’ in the moral community), that some kind of imbalance of forces in the moral community would thereby obtain (one exactly of the kind that reciprocity is supposed to prevent: *no rights without duties*).¹⁰⁶ But

¹⁰⁶ This is of course not a very plausible thought, which I would like to avoid ascribing to anyone and discussion of which in the main text I therefore want to avoid as much as possible. Nobody, I hope, thinks it unfair, say, to welcome a certain member of the family to participate with beneficiary status who is utterly incapable of carrying a moral burden herself. Nobody would call her a ‘free-rider’ in any sense suggesting that ‘reciprocity’ has been violated, and quite properly so. The name for this station that individuals occupy in

even if the story is not meant to exclude non-human animals (from the scope of practical love) on the ground that it would be ‘unfair’ to allow an enormous amount of ‘free-riding’ (as it were) but on some other, more recognisably moral ground (e.g. on the basis of the fact that non-human animals don’t participate in the Moral Law which makes human beings special in their own eyes and even, if you will, in the eyes of God), I cannot buy it.

My basic tendency in relation to this issue is in fact to think that *the more* sophisticated the story of human reciprocity (and human moral uniqueness in general) becomes, *the more* it favours recognising non-human animals as falling under the *scope* of practical love on a par with human beings. I find it particularly odd to see human beings claiming to uniquely be representatives of the All-loving on this Earth, to be in special co-operation with the All-loving Himself as those who are called to be his servants and carry out his will as a matter of moral duty, and then pointing to this peculiarity of theirs (this remarkable fact about their existence) as justifying the exclusion of non-human animals from the scope of love.¹⁰⁷ What I see people doing here (those who really mean to be excluding non-human animals from the scope of practical love on the basis of *human moral superiority*) is deceptively introducing an irrelevant factor, one that nobody would miss if it weren’t introduced and made to look relevant (a factor that is of course *otherwise* very relevant and therefore has some potential to be sold as relevant everywhere, even where it is not), with the sole purpose of blocking a conclusion that presents itself as unavoidable given the nature of our concepts and the whole of our *moral experience*: morality essentially involves being benevolent or kind (in the sense of ‘practical love’) to *all sentient beings* (i.e. to *all natural objects of benevolence*).¹⁰⁸

the *moral community* (who partake of the ‘benefits’ without carrying any moral burdens) is not ‘free-riders’ but *passive citizens* (see subsection 6.2 below).

¹⁰⁷ If there is a God whose image we bear, surely he loves all animals quite independently of their moral nobility, and all the rest just follows in the theological picture that we might like to paint on the basis of our ethical theory: our *Imago Dei* means that we are obliged to care for all sentient beings as the objects of God’s love, as the intended beneficiaries of the love that is behind the very creation of the world (see 5.4 below).

¹⁰⁸ By ‘natural objects of benevolence’ I don’t mean objects to which we ‘naturally’ (non-morally) direct our benevolence, but simply – what relates to the point being made above in this paragraph – objects of the kind that fall under the scope of benevolence given the definition of benevolence *merely as benevolence* (without yet

I want the benevolence of others, therefore I ought to be benevolent to every other. The real ‘derivation’ that Kant is proposing here for the existence of a moral requirement of universal benevolence or universal practical love (i.e. the one that we can make proper sense of) goes as follows: I (as a being *capable of happiness*) have an interest in the benevolence of others, therefore I ought to be benevolent to all those who (because they are in the same *initial* condition as me, i.e. all beings capable of happiness) have an interest in *my benevolence*. I can agree that our interest in the benevolence of others is always – given recognition of the nature of things – an interest in the benevolence of *human beings*. But the interest is not itself one belonging to our human (morally rational) nature, of course, since it is completely parasitic on our (distinctively animal) interest in our own happiness: it’s only because we desire our own happiness that we come to desire (in the sense of a felicitic desire) the benevolence of others. Accordingly, there is an *input* (matter) in our thinking referencing our animal nature (our sensibility and our natural interest in our own happiness) as its source, and you would expect that our final maxim of generalised benevolence should embody recognition of this particular provenance (i.e. the provenance of the relevant *matter*, not of course of the obligatoriness) by extending its *scope* to all animals.

Some people will of course try to make me look as putting the emphasis in the wrong place and drawing a hasty conclusion from the fact that ‘animal nature’ or ‘sensibility’ (our capacity for happiness in particular) figures *somewhere* in our moral thinking. I will let them make their argument. The truth is that it is them, and not me, who is in the business of placing special emphases. Let me put the emphasis in the fact that the ‘benevolence of others’ which one can initially desire is always the benevolence of *human beings*: in my picture, this is respected in the final, moral binding, maxim of action in the fact this final maxim is the maxim of *a human being*, a maxim that I also implicitly prescribe (if only in my mind) *for every human being*. Let me now put the emphasis on the input referencing our ‘animal nature’: in my picture, this is respected in the final, morally binding, maxim of action in the fact that the *scope* of the moral concern that this maxim

thinking of it as a moral requirement). A human being is a ‘natural object’ of benevolence because it is possible – metaphysically speaking – to be benevolent to him.

constitutes extends to all animals *as objects*: the benevolence that morality requires of us (human beings) is – as I have it – benevolence *to animals (sentient beings)*.

4.2 Kant and the Notion of Self-Givingness

This is the end of my argument for non-human animals being proper objects of our duty of *practical love*, which has also given us a basic way of ‘justifying’ (explaining to ourselves – making sense of) the existence of the duty itself. I will keep doing work that contributes to the latter kind of understanding, so I really don’t need to stop and make the point quite explicitly.

But have I done enough to justify the *for-the-sake-of relation* that grounds the *directedness* of our ‘duties to animals’? Have I said enough to the effect that morality involves acting *for the benefit of animals*? I think my arguments so far speak clearly to that effect, but I can still say more. In fact, there is more in Kant that I can bring to bear in support of the idea that morality essentially involves an element of self-givingness, i.e. of *acting for the happiness of subjects*.¹⁰⁹ Let me quote his longer rendition of the derivation of the maxim of universal benevolence (MM 6:393):

The reason that it is a duty to be beneficent is this: since our self-love cannot be separated from our need to be loved (helped in case of need) by others as well, we therefore make ourselves an end for others; and the only way this maxim can be binding is through its qualification as a universal law, hence through our will to make others our ends as well. The happiness of others is therefore an end that is also a duty.¹¹⁰

¹⁰⁹ Since I take it that non-human animals are subjects for whose happiness we are called to act in self-givingness, the resources that I take from Kant in support of the general idea of self-givingness also support the thought that morality essentially involves acting for the happiness of non-human animals.

¹¹⁰ Let us not be distracted (in the main text, I mean) by the fact that Kant is using the word ‘beneficent’ as opposed to ‘benevolent’. He thinks of ‘beneficence’ as an *active* kind of benevolence – in fact calling it “active benevolence” (MM 6:452) – which needs to be distinguished from a passive kind (the act of merely *wishing* someone well). Now, one of the reasons Kant distinguishes these two kinds is to make the point that whereas in ‘passive benevolence’ we can “be *equally* benevolent to everyone” (MM 6:452), in active benevolence we cannot. More precisely, he thinks that we can allow for degrees of ‘practical benevolence’ without violating

Saying, as Kant does, that we “make ourselves an end for others” merely through “the need to be loved” (MM 6:393) is to admit that there is an input in our moral reasoning referencing our animal nature: after all, it’s only animals (sentient beings) *as such* who have a *need to be loved*. This reinforces the thoughts that I already developed, but I think we can go even further. What we have so far is a need (for love) the recognition of which, in conjunction with the formal principle of reason, generates the idea that we ought morally to *make others our ends*: “the only way this maxim [the maxim of benevolence] can be binding is through its qualification as a universal law, hence through our will to make others our ends as well” (MM 6:393).¹¹¹ But making ourselves an end for others (as in the initial step) is to *make others a means for us*; and making others our ends (as Kant says we ought to do when it comes to the final, universal, maxim of benevolence that is morally binding) is to make oneself into *a means for them*. Some people may find this strange (because rumours and myths abound in philosophy), but here we are reading Kant saying that in morality we make ourselves into *a means for the happiness of others*. His ‘argument’ goes: I need others to be there (as a means) *for my happiness*, therefore (on pains of inconsistency) I ought to be there (as a means) *for the happiness of others*.

We have been told that ‘Kantian morality’ is done for duty’s sake and duty alone, and therefore not for the happiness of subjects. I don’t deny that Kant himself said this multiple times, nor do I want to deny that there is a construal of this particular dictum that makes it true: the only *obedience* that is of itself part of morality is *obedience to the moral law*

the universality of the maxim of generalised benevolence. I don’t stand by these claims (at least as they are stated) because I don’t allow for degrees of benevolence in ‘practical love’, nor am I even clear that this is what Kant really wants to say given that it might be more plausible to talk about ‘degrees of attention in benevolence’ – in virtue of some idea of the ‘distribution of responsibilities’ among human beings, perhaps in interaction with some notion of there being ‘patients’ who in various ways *need* more direct attention than others, etc. – without having to talk about ‘degrees of benevolence’. I am not necessarily registering my final view here, only trying to read Kant’s larger claims in a plausible way. None of that matters here, however. What I really care about (in the context of my argument in the main text) is that Kant is talking about an active kind of benevolence (i.e. about what I am calling *practical love*, or at least about a part of it since we also need to introduce *non-maleficence* in the picture as well).

¹¹¹ The square brackets are mine, to indicate by name the maxim Kant is talking about.

within. In terms of what authority it is that we have to obey in morality, animals are not the answer, happiness is not the answer, and only *reason* is the answer: all obedience in morality is obedience to the moral law of our own reason. However, this point doesn't detract from the fact that Kant is at the same time committed to there being a perfectly respectable sense of the notion of *acting for the happiness of subjects* (which we identify as essentially belonging to morality, and which therefore has to be consistent with the motive of duty being the sole genuine moral motive, which now we are only describing further) and therefore of these subjects themselves figuring as *intended beneficiaries of morality*.

What is crucial from the perspective of Kantian philosophy is that we do that which constitutes our duty simply *because it is our duty*; and even though I see how talking about happiness might be perceived as introducing a new 'because' that is potentially in tension with the first one (a potential threat to the 'purity of morality', a 'reason of the wrong kind' that seems to threaten the 'reason of the right kind' already in place), the truth is that it just doesn't do that. There is no 'because' here that could compete with the authority of reason, no extra moral-normative property being posited as *ground* of the relevant duty, much less an additional incentive of self-love that needs to come as a supplement for the moral incentive (which I agree has to be 'pure' in the sense of 'independently effective'); what is being introduced is only a *for* connecting us with the *intended beneficiaries* of the relevant duty, that's all. There is no contradiction in saying that we *ought morally* to do something *for someone's happiness*, and so it seems to me that I can say: yes, I am doing something 'simply because I ought to', *simply because I ought to do something for someone's (i.e. everyone's) happiness*. Here, as before, no contradiction is apparent. Through the notion of 'practical love' that features as a fundamental moral requirement, a special and irreducible bond is posited between the moral being and all animals as the *intended beneficiaries* of her conduct, a bond (called 'practical love') involving the notion of *acting for the happiness of subjects*. Through this notion of practical love as a fundamental (i.e. both essential to morality and universally pervasive of it) moral requirement, our most general 'moral slogan' becomes: *from duty, for everyone's happiness*.¹¹²

¹¹² The same 'slogan' was already introduced by Jacob Lautrup Kristensen in his 2015 doctoral work: *From Duty, For Happiness – Authority and Value in Kant's Theory of Practical Reason*. The 'slogan' is used in that work

5. Respect for Animals

Having prepared the ground as much as I have, I should by now be able to explain why the notion of ‘practical love’ entails – as I intend it – the notion of ‘respect for animals (sentient beings) as such’. But I cannot get to that without giving a last chance to the ‘Kantian indirect view’, if only to try to persuade us that there is some important truth that it contains and we haven’t yet seen (for we have seen a few already). Here we run the risk of taking a few steps back, since I will have to call into question (if only for the sake of the argument) the idea that ‘practical love’ – as developed in the preceding section – entails ‘duties to’ and therefore ‘respect for’ non-human animals; however, this just seems necessary for me to try to do full justice to Kant’s ‘indirect view’.

Thus, someone may grant that we ought morally to be benevolent to animals (even as a matter of ‘practical love’, i.e. benevolence *from duty*) and still deny that we ought to do anything *for the sake of the animals themselves*, thus blocking the notion of ‘duties to animals’, and therefore ‘respect’, and therefore ‘dignity’, etc. And so we are transported back to Kant’s idea of what it means to ‘have a duty to someone’ and the recognition that non-human animals don’t fit the relevant bill. Kant’s claim, remember, is that ‘having a duty to someone’ means ‘being constrained by that someone’s will’ (MM 6:442), and what it takes for someone’s will to constraint morally is for it to have the *moral authority* to do so, first, and second to be able to address its claim or demand to a suitably responsive will (i.e. to be an object of experience for the subject that is meant to be bound), all of which is beyond the powers of *non-human* animals because they are assumed to lack the original authority that would allow them to present themselves to us as ‘making moral demands’.

Now, I am omitting the problem that if we take Kant’s definition of what it means to ‘have a duty to someone’ *as it reads*, i.e. as *literally* requiring that the ‘obligor’ (the

as an abstraction for a general argument concerning Kant’s dualism between reason and sensibility in philosophy (especially moral philosophy) addressing some of the themes that I have myself been discussing. This work is similar to mine in that they are both trying to ‘re-think’ the moral significance of happiness / hedonic well-being (in something like the way Utilitarians typically have it) in Kant’s moral philosophy. This work is recommended for readers interested in a wider – and in many ways deeper – examination of Kant’s philosophy which is also more conversant with current Kantian scholarship.

obligating subject) have the power to impose a moral constraint on the obligee's will (the will of the subject putatively being placed under obligation), then – given Kantian autonomy – the conclusion follows that *a human being can only have duties (i.e. ethical duties) to himself*. No one apart from myself can make the moral law and its commands authoritative upon my will: not the animals, not my fellow humans, only I alone as a morally autonomous being. I am omitting this 'problem', I say, in the sense of assuming that nobody in this discussion is meaning to step over it. I am taking it for granted that advocates of the 'indirect view' (who want to allow for 'duties to human beings in general' while denying 'duties to animals') are arguing on the basis of some account of 'bipolar obligations' (a.k.a. 'directed duties', 'duties to someone') that *doesn't* require 'other human beings' having the power to *directly (or even literally)* place the individual under obligation. This is a safe assumption to make given the most prominent stories that have been told in this area, including Stephen Darwall's.¹¹³

5.1 The Contractualist Challenge

Granted that human beings cannot be *morally* obligated except through their own self-obligating power (i.e. their moral autonomy), it remains the case that human beings still figure, for every other human being, as having a shared *second-personal authority* to 'make claims' and 'address demands' (Darwall, 2006). We might agree to describe human beings as 'obligors' and 'obligees' with respect to each other on the basis of the fact that every human being's moral prescriptions are necessarily *universal* (i.e. given that human beings always implicitly prescribe *for all human beings*); or maybe we only agree to describe human beings as 'co-legislators' in the moral community in a way that doesn't entail their actually being *obligors* and *obligees* with respect to each other but only with respect to *themselves* (in which case the act of co-legislation would have to be understood as the moral agent putting

¹¹³ Darwall's account of bipolar obligations appears in *Bipolar Obligation* (2012). However, he puts this particular point better in *The Second-Person Standpoint*, which he does in the name of "Pufendorf's Point" (Darwall 2006, p.112): "in holding people responsible, we are committed to the assumption that they can hold themselves responsible by self-addressed demands from a perspective that we and they share." (Darwall 2006, p.112)

herself self *alone* under obligation, yet on the basis of prescriptions that are *universal*), leaving it open whether there is a more adequate name to for the role that they play in this particular inter-personal moral interaction (I mean, apart from *co-legislators*, on which we can safely agree).

Whatever we decide is the best name to use here in addition to ‘co-legislators’, what we are asked to acknowledge (before even doing any detailed conceptualisation) is that there is a special human-to-human moral relation whereby human beings stand to each other as *moral equals*, as being under shared moral standards (‘common laws’) who can *on that basis* ‘hold one another to mutual obligation’ (Darwall, 2006, p.119). This is the sense of ‘reciprocity’ that Kantian (Contractualist) philosophy is supposed to emphasise: ‘reciprocal recognition’ between human beings as *moral equals* (as beings with equal moral authority in virtue of their equal participation in the Moral Law).

Now, it should be noted that this concept of ‘reciprocity’ is closer to the ‘Christian’ idea of relating to human beings as *image-bearers of God* (as co-representatives of God on Earth) than it is to some of the other notions that have come up in relation the scope of the duty of generalised benevolence or ‘practical love’, e.g. the notion that ‘giving’ (in benevolence) is ‘obligatory’ only on condition that the ‘receiving party’ also have the duty to ‘give’ *in reciprocity*. But remember that here we are assuming benevolence to all animals to be a genuine duty, and that the story is only trying to undermine the claim that there exist *duties to animals*, duties that we owe it to animals themselves, which it does by appealing to the notion of *co-legislation* as one where only human beings can figure as direct participants: if only human beings participate in relations of co-legislation, and if it is the case that the duty to participate in these relations comprises all interpersonal duties so understood (i.e. all duties to others, whatever their content may be), then the conclusion follows that there are no duties to non-human animals.

If we want to generalise the point and include (Kantian) duties to self in the picture, we can say: One can have duties to someone only in virtue of *self-legislation* (duties to self) or in virtue of *co-legislation* (duties to other human beings). In relation to non-human

animals, we therefore say: we owe it to ourselves and to each other (or perhaps to human beings or humanity in general) to be benevolent and care for their happiness.¹¹⁴

5.2 A Contractualist Reply to the Contractualist Challenge

Having added more distinctively Contractualist (Kantian) elements to the background story that may still be told in support of the ‘indirect view’ in the context of ‘animal ethics’, we can ask: is it plausible to deny the notion of duties to animals (i.e. to non-human animals)? And the answer still seems to be *no*: emphasising the facts of ‘human moral uniqueness’ in all their metaphysical grandeur, even making ‘reciprocity’ a spiritual bond between human beings where other animals enjoy no participatory status (because they lack the authority to do so and therefore cannot count as our moral equals), is not going to help the ‘indirect view’ gain any relative plausibility as against the one that I am proposing in this work, since I am assuming those same facts: the same ‘human moral uniqueness’ (see **Part I**) and the same positive facts about distinctively human-to-human moral relationality (the reference to God in particular being only metaphorical as far as I am concerned, but the spiritual bond between human beings something very real). There is such a thing as reciprocal recognition between human beings as moral equals, a relation whereby we stand together as co-workers in the moral community (as perfect equals in that regard), and yet there are

¹¹⁴ Two accounts that seem relevant here are those offered by John Rawls in *Theory of Justice* (Rawls, 1999, p. 448) and Thomas Scanlon in *What We Owe to Each Other* (Scanlon, 2000, pp.185-186). These are two largely ‘negative’ views in ‘animal ethics’. But ‘animal rights’ are not exactly absent in the history of what Rawls would call the ‘contract doctrine’, since we find them, for instance, in Jean-Jacques Rousseau (1987, pp.35-36) [putting an emphasis on key terms]: “An end can also be made to the ancient disputes regarding the *participation of animals in the natural law*. For it is clear that, lacking intelligence and liberty, they cannot recognize this law; but since they share to some extent in our nature by virtue of the *sentient quality* with which they are endowed, one will judge that they should also *participate in natural right*, and that man is subject to some sort of *duties toward them*. It seems, in effect, that if I am obliged not to do any harm to my fellow man, it is *less because he is a rational being than because he is a sentient being*: a quality that, since it is common to both animals and men, should at least give the former the *right* not to be needlessly mistreated by the latter”. Contractualist (Rawlsian) arguments for the moral standing of sentient beings in general have already been produced, one prominent example being the one offered by Mark Rowlands in his book *Animal Rights* (2009).

duties to non-human animals. Let me briefly discuss an example (often treated in the literature) that I think helps to illustrate my point.

Saying that we only have duties to human beings, and that ‘duties to animals’ are on this basis out of the question, is of course to say that only human beings can be *wronged* when mistreated. Thus, if I intentionally step on someone’s (i.e. a human being’s) foot without a good enough justification, I *wrong him* because (as a human being) he has the authority to make claims on me, including the claim that I don’t step on his foot without proper justification. If, however, I intentionally step hard on a dog’s tail for entertainment, I am certainly mistreating the dog and acting immorally; but since the dog lacks – as Darwall would put it – ‘second-personal authority’ (for he lacks the moral lights that someone would need to be able to generate moral claims and demands), it turns out that *I don’t wrong him*. Myself and other human beings perhaps, whom I let down as fellow moral authorities, but not the morally blind creature in front of me, whom I only *harm*. If the animal looks at me, it is perhaps with the *fear* that I am going to do something more severe to him; he is not invoking (at least not intentionally) the sacred law that addresses me and my fellow humans and on the basis of which ‘we make claims of each other’; therefore – I am supposed to conclude according to dominant ‘Contractualist’ doctrine – the animal in front of me has no ‘rights’, no duties *directed to him*. I ought to be kind to him, surely, but I don’t *owe it to him*.

But it appears to me manifest that if I (as a human being) have a special authority to issue the demand that nobody should step on my foot without good enough justification (to continue with our example), then I have *the very same authority* to demand that nobody should step on the dog’s tail, and I can use the language of ‘respect’ and ‘rights’ and ‘personhood’ in relation to the subject whose interest I mean to protect through the issuing of the demand (the intended beneficiary), *in the one case just as much as in the other*. I can say: respect my right! And I can say: *respect his right!* This is of course the idea of ‘giving animals a voice’, which of course only means a ‘moral voice’, which in turn only means: representation of their interests (as *owed* to them in their condition as intended beneficiaries of morality) *vis-à-vis* human beings in general as co-legislators and – I would

now like to add – *co-workers* in the moral community.¹¹⁵ This idea of ‘giving animals a voice’ is itself a piece of Contractualist doctrine, the upshot of the interaction between the ‘Kantian’ idea of co-legislation and the *Utilitarian* or *Agapist* idea of ‘practical love’ for all sentient beings.¹¹⁶ This simple story seems to me to reveal (in a very powerful way) that inter-personal human-to-human moral relations are in fact more complex and at the same time more interesting than the bare ‘Kantian’ idea of co-legislation – if spelled out in isolation – could possibly convey: co-legislation is as such irreducible to any other putative moral relation, but it seems to be nothing more than an aspect (an fundamental and very important aspect) of our more general relation as *co-workers* in the moral community, as *co-guardians* in a moral community essentially including all sentient beings as having ‘beneficiary status’.

There are two more positive cases that I can make in support in support of the notion of ‘respect for animals’, i.e. for the notion of ‘practical love’ for animals (which at this point we all agree on) being one according to which *we owe it to animals*, i.e. for our ‘practical love for animals’ being a genuine *duty to animals* (a duty to *respect animals*). The first case is meant to show that theological ideas which could be appealed to in support of the ‘indirect view’ are actually themselves *better understood* (as narratives that could follow from a genuine concept of morality already in place) in light of a moral theory allowing for the notions of ‘duties to animals’ and ‘respect for animals’. The second is to make sense of the notion of ‘animal rights’, which is very often thought of as having no room in the

¹¹⁵ This ‘Contractualist’ idea is not denying that non-human animals have *their own voice* (for they obviously have one), nor is it to suggest that this voice is not of itself salient from the moral point of view (which it obviously is); in fact, it is not even to say that the voice of any particular animal out there is not in fact a *moral voice*, since that is not for moral theory to judge.

¹¹⁶ Stephen Darwall considers this kind of idea by saying that human beings may be said to act as ‘trustees’ (with respect to non-human animals) having the “authority to demand certain treatment on their behalf (perhaps also to claim certain rights, compensation, and so on, for them)” (Darwall 2006, p.29). However, Darwall also completely trivialises this idea by suggesting (through a certain literary reference) that human beings may have the authority to ‘speak for trees’ (Darwall 2006, p.29). Again, this completely trivialises the idea of ‘giving *someone* a voice’ because this idea only makes sense in relation to *persons*. A tree is not ‘someone’, a tree is not a ‘person’ or a ‘subject’, therefore it cannot have any kind of citizenship in the moral community (lacking the capacity for happiness, for instance, it cannot be said to have ‘beneficiary status’).

Kantian picture. Let me give you the latter first and the former second (each taking a separate subsection).¹¹⁷

5.3 *Animal Rights*¹¹⁸

The ‘Kantian view’ (on ‘rights’) has often been portrayed as saying that: since nothing about non-human animals themselves can serve as the suitable ground of ‘rights’, therefore non-human animals lack rights. That since – to put it more generally and more to the point – felicitic interests as such (the capacity for happiness or well-being) cannot be the ground of any rights, therefore individuals cannot come out as suitable right-holders on account of merely possessing interests of this kind. Human beings, who have rights, don’t have them because they have a capacity for happiness or well-being (or in any case ‘interests’ of some relevant kind),¹¹⁹ but for some other reason. But what is that reason exactly? Well, the thought suggests itself naturally that individuals only acquire rights *through themselves possessing duties*. If correct, this thought would offer a clean explanation of why – as most Kantians would have it – all human beings, and no non-human animals, have rights. But let us put this thought to the side for the moment, since I don’t think this is what the ‘Kantian view’ is really trying to say.

¹¹⁷ You can see me as closing my main reply to the ‘Contractualist Challenge’ at this point. It only takes showing that there is more to morality than people seem to assume who put forward that challenge. However, my reply actually continues in the following three subsections, which are meant to *give support* to the basic point that has been made in this section.

¹¹⁸ The concept of ‘animal rights’ is not identical with the concept of ‘the rights of non-human animals’: *animal rights* are the rights that all animals possess *simply ‘as animals’* (i.e. simply as beings *capable of happiness*), which entails that human beings – who are animals in this sense no less than pigs and dogs are – have ‘animal rights’ themselves. Also to be clear, there is a place in the view that I am proposing for ‘distinctively human rights’, rights that only human beings have, but there is more to ‘the rights of human beings’ than *that*: there is also – as I am arguing here – ‘the *animal rights* of human beings’. By the way, I conceive of ‘animal rights’ as always having the sense of *beneficiary rights*, rights that individuals have to participate in the moral community with ‘beneficiary status’.

¹¹⁹ Except perhaps ‘irreducibly moral interests’.

What I think the ‘Kantian view’ is trying to say, and what I myself take it to say in the version of it that I endorse, is that individuals acquire ‘rights’ only in virtue of having duties directed to them, in virtue of someone having *duties to them*. If this is so, then the actual commitment of the ‘Kantian view’ in relation to non-human animals is the following: for non-human animals to have ‘rights’, it has to be shown that we have duties to them. And hence of course we realise that we have already been giving an argument for ‘animal rights’ as they are understood in the ‘Kantian view’, since we have been arguing for those duties which are meant to ground those rights. The idea of ‘animal rights’ has as much room in Kantian ethics as the notion of ‘duties to animals’ does, and they also stand or fall together because the one is a mere corollary of the other.

Defending the notion of ‘animal rights’ becomes difficult when one tries to make sense of the thought (that many people associate with the so-called ‘interest theory of rights’, which is meant to stand as an alternative to the ‘Kantian view’) that all animals or sentient beings have rights (i.e. ‘beneficiary rights’ in particular) *merely* on account of their capacity for happiness. Well, I guess we only have to be clear on what we mean, and what we don’t mean, when we make that claim. By now it is clear that we don’t reference the capacity for happiness in order to locate the *ground* (the ultimate metaphysical ground) of ‘beneficiary rights’, which is now understood to lie in our duty of practical love to animals. No, what we mean to be doing in saying that ‘animals have rights (of the relevant kind) *merely* on account of their capacity for happiness’ is simply that they fall under the *scope* of the relevant kind of protection because they have a capacity for happiness, because the principle that enjoins this protection (practical love) is essentially directed to all animals *as intended beneficiaries*. The ‘merely’ qualifier which we use to emphasise the point is again not to suggest that the referenced capacity is ‘all that it takes’ to generate the existence of ‘beneficiary rights’; rather, it is only to note that individuals count as right-holders (holders of ‘beneficiary rights’ in this case) independently from themselves having or lacking any other capacity (including the capacity for morality, which may be otherwise relevant *but not here*).

It has always been a major objection to the ‘Kantian view’ in the ‘theory of rights’ that it is far too restrictive in terms of its scope of protection, but this deficiency can be amended in a principled way through clarification of our duty of practical love in its

directedness to animals in general *as intended beneficiaries* (with the qualification, of course, that ‘animal rights’ / ‘beneficiary rights’ are not meant to be exhaustive of the theory of rights).

5.4 *Animals (Human and Non-Human) and Their Place in Creation*

If the traditional Kantian story (purportedly supporting the ‘indirect view’) can be given ‘theological support’ (in the sense of their being a good story to tell on the basis of a moral theory already in place to begin with), mine has every right. I can, for instance, accept the notion that human beings are uniquely *image-bearers of God*, and therefore I can also accept the idea of us having a duty to relate to every other human being as a *co-image-bearer of God*, as a co-worker in the moral community for instance. All of this I can accept. My point, then, is that this story requires a few more elements added, even for the sake of understanding what it means for us as individuals to be and act as bearers of God’s image.

Thus, to put my cards on the table, I submit that God did not create the world for the sake of human beings exercising their moral powers, but *first and foremost* for the sake of *sentient beings (his children, human and non-human alike) enjoying life*. God created sentient beings and the world as a whole (his children and a home for them to inhabit) for the sake of having *someone to love*, and he created human beings as morally rational beings (his representatives) for the sake of having *someone through whom to love in the most thorough way*. Given the reality of this spiritual idea (which we find present in the phenomenology of self-givingness, i.e. a basic sense of existing for those subjects who are the beneficiaries of one’s love), it follows that a human being’s bearing of God’s image cannot be exhausted by her appearing in front of other human beings as having moral authority; in fact, it consists first and foremost (in terms of the point of her mission in this life) in being there to see to the benefit of all sentient beings, human and non-human alike. We are God’s hand in the world, and in this sense his subjects, his ‘servants’ if you will, standing alongside his other servants (all other human beings) as *moral equals*. However, we are also, and by the same token, the ‘servants’ of all animals: we are *guardians* in a moral community where all animals (all of God’s children, none of whom God forgets) participate with beneficiary status. Again, we (human beings) are uniquely God’s representatives on Earth, but surely this

entails being *God's hand* in the world, and the thought suggests itself naturally (at least to me, of course, since this is the kind of 'theological narrative' that I would tell on the basis of my own ethical theory) that we ought to be *God's hand by the side of the creatures* (i.e. we are here to see to be benefit of God's children, all of whom are the intended beneficiaries of his love quite independently of their moral nobility).

5.5 Summary and Conclusion

To recapitulate. What advocates of the 'indirect view' have been doing is taking the moral duties in question (our duties to treat animals well) and portraying them as involving *only one* fundamental moral relation, that between the acting subject and her own self-legislation, between her capacity to form practical intentions and the deliverances of her own moral reason, thus saying that these duties are ones that we owe to ourselves as morally rational beings, and not to animals. Now, since it is part of their project to also make sense of *duties to human beings in general*, they allow for *co-legislation* (what is obviously a *human-to-human* relation) as the only fundamental moral *inter-personal* relation, the only fundamental moral relation apart from the already stated intra-personal relation of self-legislation. And then what they do is portray their view as very simple and elegant, construing all fundamental moral relations as just various ways of *honouring the moral law* and having to introduce nothing more by way of the *matter*, or the *point*, or even the *form* of morality. It says: in light of the idea that all human beings are under the moral law (which puts them under obligation directly), and in light of the fact that they all share this law (which puts them all on a plane of *moral equality*), we can account for duties *to human beings* in general, 'duties to self' and 'duties to other human beings', end of the fundamental philosophical story. It is true that we have to care about non-human animals and their happiness (as a matter of practical love even), but they only figure as part of the 'content' of morality and have nothing to do with the 'form', which is what matters most to philosophy: things like 'unconditional necessity' and 'universality' and 'justifiability' and 'accountability', etc. In any case, they say, we can be benevolent to animals as a matter of moral duty without ever having to posit a single 'duty to animals'.

The reason I don't buy this orthodox Kantian story (orthodox, I say, yet by now somewhat improved with respect to the one that Kant himself gave us, which goes as far as to reject practical love in relation to non-human animals, in fact regarding these animals very much as commodities) is because I take there to be more to the fundamentals of human moral relationality than just *self-legislation* and *co-legislation*. Now, since the orthodox Kantian picture (as represented by the 'indirect view' in 'animal ethics') presents itself as the complete philosophical story on fundamental moral relations, then we have to conclude that this picture is unduly reductionistic (a picture that unduly explains away certain fundamental aspects of morality).¹²⁰ Morality (the whole of morality, morality as

¹²⁰ Here ends the argument for the orthodox Kantian picture (as represented by the 'indirect view' in 'animal ethics') being guilty of a peculiar *false reductionism* of its own: a picture that unduly explains away our 'duties to animals' and everything that comes with them (including the requirement of practical love in its full signification and dimension, a requirement that calls the moral agent to make herself into an instrument for the benefit of animals in general as an act of self-givingness). Let me add to this case (which I do here because I think it would ruin the cohesion of the argument in the main text) by pointing out some dialectical facts. Advocates of the 'indirect view' and I can in one voice reject the *false reductionism* that some people (some Utilitarians, for instance) have fallen into when they point to happiness itself as *moral ground*, and we certainly can to a large extent diagnose this error in the same way. The reason people are tempted to go in that direction may be that happiness presents itself as having *intrinsic value*, but this value is of itself *non-moral* and will never (however you take it) amount to the morality of actions. The idea is also tempting because happiness (the happiness of all subjects in this case) presents itself as morally incumbent upon us in some very fundamental way, but this moral incumbency (its 'moral-practical value', its '*ought-morally-to-be-pursuedness*' if you will) attaches to it only as a *represented end* (i.e. not in itself, precisely because felicitic outcomes are not themselves *part of morality*) and only in virtue of there being a principle in place demanding that we pursue it, a principle of practical love (in this case) addressing our own will *directly*. Happiness is not as such part of morality, nor is its pursuit morally required just because it is at stake. Ultimately, however, I don't think that advocates of the 'indirect view' have all the resources that are needed to offer a real and satisfactory diagnosis of the mistake that some Utilitarians have been making (their peculiar kind of false reductionism). They can say as much as I just registered, but they cannot really explain why on Earth there is a view in *moral philosophy* (defended by human beings doing a more or less honest and dedicated effort, in many cases doing great work in all kinds of areas) reducing everything *moral* completely to the felicitic. Knowing them very well, I bet the reason they are doing this (i.e. the reason for their philosophical mistake, at least in many cases) is not simply that happiness presents itself as having 'intrinsic value', or even because everyone's happiness presents itself as morally incumbent in some fundamental way, but because they think of it as that which *ultimately matters*,

such and therefore *every bit of it*) has the sense of *self-givingness*, of something that one does *for the benefit of all subjects*, which in my view means *for the happiness of sentient beings*. Morality essentially involves a duty called ‘practical love’ that, because of its specificity (e.g. the specificity of its direct object: sentient beings in general as intended beneficiaries), the *sense* that it brings to morality (self-givingness), and the fact that it has its origin in the moral being as such, has to be counted among the fundamental moral relations that should feature in the very structure of our ethical theory. I am even claiming that this particular relation is ‘universally pervasive of morality’ in the sense that it features as an essential aspect of *every moral duty that we have*.

A central reason why we have to reject the orthodox Kantian picture (which we have been talking about from the very start) is that it fails to recognise all animals as persons. In virtue of the ‘duty’ of practical love (which is actually more like an aspect of every moral duty that we have), morality has the sense of *self-givingness*, of doing things *for the benefit of all sentient beings*. But this self-givingness that morality involves only makes sense in relation to *persons*, therefore *all animals are persons*. All animals have a right to participate as citizens in the moral community with beneficiary status (i.e. only animals have duties directed to them as intended beneficiaries); but only persons have rights, therefore all animals are persons. All animals are to be recognition-respected in their beneficiary status in the moral community; but only persons can be recognition-respected, therefore all animals are persons.

which I argue is true – *in some sense* – even in the Kantian picture properly understood (as enjoining self-givingness in practical love). Thinking of happiness (everyone’s happiness) as that which ‘ultimately matters’ (which I argue is a thought difficult to spell out in its moral signification and which – given dominant moral doctrine – is often taken as an attack on morality), they think that they can build their moral theory with the felicific as their ‘moral starting point’, thus portraying it as ‘the first moral reality’ or ‘the basic moral fact’ or as containing ‘the most fundamental moral distinction’ (*viz.* the value distinction – which they assume is a ‘moral distinction’ already – between pleasure and pain). This explains the error as a philosophically naïve mistake that doesn’t necessarily involve any kind of genuine attack on morality (although it can mislead people a great deal and doesn’t make the philosophical gods any proud; and it may on occasion be a genuine attack, I also have to admit). This is a nuance of diagnosis that advocates of the ‘indirect view’ are incapable of conveying, since they seem to be claiming not to see the most important *truth* behind the mistake in unduly reductionistic versions of the ‘direct duty view’.

We have often been told that including in morality any reference to the happiness of subjects is a threat to the purity of morality. But this cannot be right. What is a threat to morality is to include the *wrong* reference to happiness (trying to base morality on the pursuit of one's own happiness in self-love being the one that Kant combated the most). In fact, if you think about it, there are references to *morality* that constitute a serious threat to the purity of morality. It is real threat to morality, I submit, to give morality for answer to the question (the fundamental moral question) as to what capacity it is that fixes the domain of *intended beneficiaries of morality*. A massive rationalisation that ends up robbing the majority of persons of their personhood and their dignity and their rights (their 'beneficiary status' in this case) and which at the same time, and by the same token, disenfranchises essential aspects of our humanity (e.g. our duties to non-human animals) is a real threat to morality. The only thing that guarantees the purity of morality is always giving the right answer to moral questions, whatever that may be.

6. Animals, Babies, Persons, and Human Beings

I am not here joining the debate surrounding the issue of 'marginal cases'. Let me express myself better: I am addressing some of the issues that are normally talked about under the rubric just mentioned, but I am rejecting the terms of the debate as most philosophers have been having it. I have always felt there is something wrong going on, some unfairness from the start in the very way that concepts are employed, and it has been some time since I realised exactly what the problem is.

6.1 *Animals, Babies, and Persons*

Talking about the moral standing or the dignity of 'animals' *vis-à-vis* that of 'persons' is an unfair deal to begin with, for all animals (all sentient beings) *are* persons. A cow, for instance, is a 'someone', not a 'mere something': she is a 'subject' or a 'person', not a 'thing' or a 'mere object'. People often compare animals to babies (i.e. human babies) to argue analogically for the moral standing of the former, but I think they should mean it *literally*. They point to the unfairness of excluding 'animals' from the domain of direct moral

considerability while at the same time including ‘babies’ and other human beings with similar mental capacities as those of the animals being excluded, which I completely understand. But there is no ‘comparing’ animals (i.e. non-human animals) with ‘babies’ *because animals are babies*. In interpreting the command (particular as it is, both in content and use of terms) ‘not to torture babies’, for instance, we are not allowed to assume that any young member of the human species (e.g. an unviable embryo) counts as a ‘baby’ in the relevant sense, nor are we allowed to assume that all members of other species fail to count as ‘babies’ in the relevant sense because they don’t belong in the ‘human species’, or even because they lack a capacity for morality, or even because they are not young member of the human or any other particular species.¹²¹ The truth (which everyone knows) is that we are called to relate to sentient non-human (non-morally rational) animals in general (however old they are and whatever biological species they may belong to) ‘as babies’, as beings with beneficiary status and no moral responsibility whatsoever, as intended beneficiaries of our own activity (‘darlings’ if you will, and certainly also ‘neighbours’ or ‘co-beneficiaries’ in the moral community) and nothing more. The ‘nothing more’ doesn’t mean that they are less important *as intended beneficiaries*, but simply that such is the only distinctive role they play in our moral thinking. Now, we need to look for a term more technical than ‘baby’ to talk about these matters,¹²² but you can see what I am getting at by

¹²¹ There are at least three possible notions of ‘baby’ that one might talk about: (i) a ‘very young’ individual (a very young member of a certain species), (ii) someone incapable of fending for herself in a minimal sense, and (iii) a ‘necessarily innocent individual’, a ‘person capable of happiness but completely oblivious to morality’. The first (age-related) sense of the term will hardly be of any moral significance, since there are ‘baby orchids’, and ‘baby orchids’ don’t count as babies for the purposes of our common injunctions not to mistreat babies. The second sense will also hardly be of any great significance in moral theory: granted that an individual (i.e. a sentient individual) incapable of fending for herself in a minimal way *needs more attention* than an individual more capable, that fact doesn’t give her any greater moral status: it certainly doesn’t make her more of an intended beneficiary of practical love (again, except in the sense – which amounts to nothing fundamental – that more direct attention should be devoted to her). The third sense of ‘baby’, however, i.e. the sense that I am pointing to in my argument, is of immense significance in moral theory because – as I will explain – it is equivalent to the category of ‘passive citizen’ which is structural of the moral community.

¹²² Which is *not* to suggest that the use of the word is somehow inappropriate for common use in relation to non-human animals. To the contrary!

appealing to a ‘technical’ version of it that equals ‘non-human animal’: (i) if ‘non-human animals’ count as ‘babies’ in the relevant sense of the term (because in this context they correspond to one and the same technical category); and (ii) if all ‘babies’ are ‘persons’ (if only ‘non-moral persons’); then, assuming that killing ‘persons’ for fun (or in any case *morally unlawfully*) constitutes murder, (iii) therefore, killing non-human animals for fun (e.g. for the taste of their flesh) constitutes murder.

That there is in fact a technical sense of a term like ‘baby’ which is *personhood-entailing* and is equivalent (not only in extension but also in intension) to the notions of ‘mere animal’, ‘non-human animal’, ‘necessarily innocent person’, ‘being cable of happiness but incapable of morality’ is attested in our ordinary moral discourse. For instance, we speak of *adoption* in relation to both human babies (and children in general) and non-human animals, and we mean it in the same sense in both cases. This is not to say that a ‘human baby’ with prospects of breaking free into a moral condition is exactly the same kind of ‘project’ that a cat (whom one adopts) is; in many ways it’s just different, but in the relevant respect (i.e. when it comes to protecting the happiness of the person for her own sake) it very much is: killing one’s cat for convenience is just as murderous as killing ‘one’s baby’ for convenience, for the cat is himself one’s baby (a necessarily morally innocent person under one’s protection). Now, the fact (which I cannot deny) that people often speak of *buying* animals while treating it as something morally innocent is not going to prove me wrong: if anything, it proves the that human beings are capable of messing up. The case is that human beings *should* be speaking about *adopting* animals, never about ‘buying them’ because animals are not ‘things to buy’.

Take also for illustration (a second piece of philosophically suggestive evidence to the effect that animals – i.e. *non-human* animals who really count as such – are properly to be regarded as ‘babies’ in the relevant technical sense) the way human beings often react when they feel disrespected in their condition as morally rational beings, for instance when they say things like: I am not an (mere) animal, or in fact: I am not a baby! The context in which phrases of this kind are ‘typically’ uttered (and of course I’m idealising here for the sake of making my point) is one where the moral agent is affirming his ‘right’ to be recognised as a morally competent agent, someone that is more than a mere intended beneficiary of morality, more than an object of practical love, but instead someone who is

here to also (and in a practical sense – from her own perspective – most importantly) play a part in the moral government of the world. But when a human being is about to be tortured for fun, she doesn't say: I am not an animal! I am not a baby!, since that would give to understand that it would be alright (or perhaps 'less bad') for her torturer to instead go after the latter, which of course is a plainly *wicked* suggestion to make. To be sure, people say things like that, they say them and appear very proud, thus also appearing very wicked. A good human being who is about to be slaughtered to have her flesh taken for someone's pleasure doesn't say: I am not an animal! I am not a baby! There she rightly says, *I am not a carrot*, since carrots are the kind of thing that exist to be eaten: 'animals' and 'babies' (i.e. all 'babies' in the technical sense of the term) *are not*. I mean, of course the person in the latter scenario would say some other thing, but I hope this manages to convey the point. What I hope you can see, of course, is how certain appeals to morality and our capacity for it are straightforwardly *wicked*, and also how there are references to the individuals in their capacity for happiness that are simply necessary to keep our philosophy in touch with moral reality and our principles in their correct scope of application.

I can prove that I am not engaging in mere wordplay here, if that is your concern. If it matters so much that we don't use the word 'baby' in relation to, say, a sentient cat (unless it be a young cat, so that we could speak of a 'baby cat', which we would then have to distinguish from 'babies in the relevant sense', i.e. babies of the kind that could potentially be *murdered* and *raped* and in other ways *wronged*, who are said to be – for some strange reason – all and all members of the human species, a.k.a. 'human babies'), then I can give you the very same point rephrased: animals (sentient non-human animals) are such that killing them for fun (or for some other less than morally justifying reason) is exactly as murderous as killing sentient human babies for fun. Well, actually this point about some acts of murder being no more murderous than other acts of murder actually transcends the issue of 'babyhood': being 'more than a baby' (in the sense of 'babyhood' that I am pointing to) doesn't make someone more of an intended beneficiary of morality, and I take it to follow from this that it is exactly as murderous to unlawfully kill 'babies' as it is to unlawfully kill adult human beings (or just 'human beings', since it is enough that someone stop being a 'baby' – in the sense of having an original moral thought – to become a 'human being' in the technical-philosophical sense of the term).

There are questions surrounding ‘abortion’ (and related issues involving the treatment of ‘human organisms’) that seem to be in the minds of many people. Let me just state some of the basic elements of my view dogmatically for the sake of formalising my claims about the nature of ‘persons’ situating them in some basic ontological and moral map. Killing cows for the taste of their flesh belongs in an entirely different category (as an act of murder, no less murderous than any other act of murder) from killing, say, an unviable human embryo for fun or for any kind of convenience, which *could* – given the details of the case, which is not to say that it is so in ‘normal cases’¹²³ – be wrong and base but could never count as ‘murder’, since murder is always the murder of *someone*, i.e. of some *subject* or *person*, not a thing. By the way, adding that the ‘human embryo’ being ‘killed for convenience’ is a viable (instead of an unviable) one *may potentially* add some complexity to our thinking on the matter, but this consideration will by no means (and for the same reason) turn the act into an act of *murder*. A ‘potential person’ (even in relation to sentience as sufficient for personhood, as I mean it to be) is not a person, but always still a thing (in Kant’s general sense of a ‘mere something’ or an ‘entity lacking moral standing’) until it be *born as a person* in the original acquisition of personhood-constituting capacities. When a sentient subject first becomes present in the world is born the ‘person’, and in the acquisition – at some point – of moral freedom on the part of this same subject (responsiveness to moral considerations) is born the ‘moral person’ or the ‘moral agent’. I was as born as an animal or a sentient being, and so already *as a person*, the moment I first *felt* a thing, and I was born *as a human being* (in the technical sense of a ‘moral person’ or a ‘moral subject’) the day I had my original *moral thought*.

6.2 *Passive Citizens*

I am claiming that animals (i.e. non-human animals) are – in terms of their moral status – something like ‘babies’, but I also have acknowledged that it would be useful to have a more

¹²³ People don’t normally ‘abort’ ‘for fun’ but because they haven’t decided to bring a person into the world. Sometimes they abort out of inconvenience, but surely they also detach themselves physically from other *things* (i.e. allowing that we are really dealing with a ‘thing’ in this case) out of inconvenience. None of that is *murder*.

technical term to employ in philosophical discourse. Fortunately, we have it. I can suggest the term ‘mere intended beneficiary’, which captures the thought of a being having ‘beneficiary status’ and nothing more, i.e. no further status or role to play in the moral community beyond simply being someone with ‘beneficiary status’. But this is not a very well-sounding term, I admit, so let me instead suggest another one that is already at use in philosophy: *passive citizens*.

This term has acquired a bit of a bad reputation due to misuses of it: due to, for instance, people like Kant saying that “domestic servants” and “women in general” are – quoting Kant himself – “passive citizens” (MM 6:314). To put it simply and without going into the details of Kant’s views here, the point is that their interests are protected by law, but they don’t themselves have a right to rule, i.e. to vote and the like, because, says Kant, they are unfit (not ‘sufficiently independent’) for that – which of course doesn’t mean they have no moral duties whatsoever in Kant’s view since they are still rational beings.¹²⁴ But let me ask: is Kant misapplying a category (or even a thousand philosophers misapplying it) reason to reject the category itself? Of course not! Or is it the case that the category is *itself* flawed and we shouldn’t be using it because of that reason? Well, that would be a very good reason,¹²⁵ but what are we going to do with ‘babies’ and ‘animals’? Are we going to posit moral duties to them to be able to protect them as our kindred? Of course we have to welcome as ‘citizens’ or ‘members of the moral community’ subjects who don’t themselves have moral duties, and thus the need for the category that I am now introducing under the name *passive citizens*: beings who fall under the protection of moral laws (with beneficiary

¹²⁴ This is delicate territory, i.e. both the subject matter and our reading of Kant’s views on the particular are delicate business, and I will avoid further responsibility here because my argument is meant to convey a very basic point in relation to the protection of subjects who – I myself assume – are oblivious to moral considerations.

¹²⁵ Kant suggests that the concept of ‘passive citizen’ “seems to contradict the concept of a citizen as such” because what makes a citizen is exactly that she is “fit to vote” (MM 6:314). In saying this, Kant was (perhaps on purpose) setting the ground for the future to call into question his own categorisation on the basis of more egalitarian ideas. But surely we can agree that there is no human (morally rational) being that is ever to be treated as a ‘passive citizen’ (human moral egalitarianism) without rejecting the category of ‘passive citizens’, which we need at our disposal to be able to protect (as citizens with beneficiary status) all beings capable of happiness who are nevertheless oblivious to morality.

status) without themselves having any moral duties. There are citizens of this kind in the moral community: they are called non-human (non-moral) animals = non-human (non-moral) persons = non-human (non-moral) subjects = non-human (non-moral) individuals, all of whom count as *persons* for moral purposes. All animals fall under the principle of *respect for persons*, since that is what they are: *persons*.

6.3 Respect for Moral Personality

You can always expect that arguments for the dignity (citizenship) of non-human animals – referencing human children for any kind of comparison – will be given the following reply: that we (‘advocates of animals’) underestimate how different human beings really are from other animals, how soon in a human being’s life a moral personality begins to express itself, thus raising the individual above the moral darkness of animality. But I don’t deny any of this. I think human beings break free into a moral condition very early in life. And if you ask me, all children (all persons who had an original moral thought) have a right to participate as equals in the moral government of the world: I summon all of them (so go ahead and tell them!) to start by giving us their original thoughts on the moral standing of animals. Let us bring ‘philosophers’ and ‘children’ of all ages and backgrounds in the room (who want to be there voluntarily, which I bet are many), to talk about the innocent subject of animals and morality. I mean, what could be more innocent than that? We always have a chance to prove just how seriously we take the moral personality of human beings, i.e. in our practical lives we always have this chance. In philosophy (which is what I am directly concerned with in this work), we do justice to it by saying: treat all human beings as *co-workers* (co-legislators and co-guardians) in the moral community, *never* as ‘things’ (which are rightly to be treated as *mere means*) OR ‘passive citizens’ / ‘mere intended beneficiaries’ / ‘babies’ / ‘non-human animals’ (who are owed unconditional practical love but nothing more, i.e. at least not in a relationally-specific way).¹²⁶

¹²⁶ The notion of ‘relational-specificity’ is developed and employed (in the context of my technical formulation of Utilitarianism and Kantian ethics) in subsection 7.2 below.

Now, since human (morally rational beings) *do* count as intended beneficiaries on account of their capacity for happiness, they are covered by what we might call ‘the formula of animality’: Respect the ‘beneficiary status’ of *all animals*! Act always with *practical love for all animals*!¹²⁷ And since the two formulas should go together, because all of these things can only be realised *together*, we can just say: respect all persons in their dignity! Now, I could extend myself here trying to unpack ‘respect for human beings as such’ and describing the way it fits together with the other ‘formula’ that I just introduced, but I will only be able to say a few things and come back to this topic in the next section (where I do more of an effort to put all the pieces together).

The idea that we ought to treat all animals with practical love is a fairly basic idea that anyone can understand, and so is the idea that fellow humans should *also* (i.e. on top of that) be treated as moral co-legislators and co-workers in the moral community, as beings who are here to participate as equals in the *moral government* of the world (through the government of their own will in its full moral relationality, of course). I also take it that the thought of the two relations just mentioned being *universally pervasive of morality* (i.e. featuring as essential aspects of all our moral duties), and the related thought that for each of them to be realised *both have to be realised together*, are not all that difficult to grasp. My favourite example should suffice for now: treating human beings *as babies* doesn’t just violate their ‘distinctively human rights’ (if I may put it like this), but also always at the same time the ‘beneficiary rights’ of babies, and their own beneficiary rights too (as well as oneself as a moral being, etc.). To speak of non-human animals in particular, surely they deserve (they have a ‘right’ to) having their guardians never be treated as babies, but as *their guardians*! To speak of human beings in particular, surely they deserve (they have a ‘right’ to) never be treated as babies, but always as (among other things) the *guardians of animals*.

This is the end of the basic story that I needed to tell about persons and the respect owed them. I will now give you a final reflection on some the central concepts that this work as a whole has been trying to elucidate, both in their meaning and moral significance.

¹²⁷ I call it ‘formula of animality’ to indicate that it’s animals (i.e. sentient beings) who are the proper objects (direct objects) of the relationally-specific attitude being enjoined by this particular principle. But of course this is not to deny that this is a principle *of humanity*, a principle characterising an essential aspect of the *human mission*.

I will devote a separate section to this and then move to my central argument in ethical theory (formulating Utilitarianism and clarifying its connection with Kantian ethics, where I return to discussing fundamental moral principles).

6.4 *Animals equals Persons*

If my arguments are correct or at least in the right direction, it follows that there must be a very fundamental way in which ‘animality’ is (as a category) much closer to ‘personhood’ than ‘humanity’ is. After all, it follows from everything I have said that (i) all and only *animals* count as ‘persons’, and (ii) humanity – as the capacity for morality – is not necessary for individuals to count as ‘persons’, *even for moral purposes*. But how exactly is ‘animality’ closer to ‘personhood’ than ‘humanity’ is? I think I am prepared to say that ‘animal’ and ‘person’ are one and the same philosophical category, and I think this is borne out in reflection about our ordinary use of concepts. When we qualify a *person* as a ‘moral person’ (which even ‘orthodox Kantians’ and people traditionally arguing on ‘similar lines’ do, seemingly without realising the door they are thereby opening), what we are doing is qualifying an *animal* as a ‘moral animal’, which is of course what we are. Or what is again the same thing: we are qualifying a *subject* as a ‘moral subject’. The categories of ‘person’, ‘animal’ and ‘subject’ are one and the same category, just different ways of naming the same reality. There is no denying that *humans beings* are in a class of their own in virtue of the specific kind of life they are called to lead, the moral life, which raises them above the rest of the world as the moral governors of the world, but by no means do they exhaust the domain of ‘persons’ because *persons equals animals*.

If someone – pretending to be raising a serious objection – says that ‘animal’ is a biological as opposed to a philosophical category, I cannot but remind him of the etymology of the sacred word he is employing: is ‘animal’, whose etymology (from the Latin *anima* meaning ‘soul’) gives us ‘animated being’ or ‘being with a soul’, a biological as opposed to a philosophical category? If – perhaps more plausibly – he questions my proposal here by suggesting that if non-human animals are ‘persons’ or ‘subjects’ it is only in a ‘psychological’ sense that has no bearing on morality because they are not (even by my own admission) ‘moral persons’ and therefore not ‘persons’ or ‘subjects’ of the relevant kind, and that

therefore I shouldn't be concluding that 'animal' and 'person' are the same category (let alone the same moral category) because the fact is that there are animals (subjects) that don't count as 'persons' for moral purposes, I can only note that this completely begs the invalidity of everything I have said in this work. We know the stories that are meant to give support to that kind of position, and I have done my best to undermine those stories.

There are two concepts here that cannot be conflated: 'moral persons', or 'moral subjects', on the one hand, and 'those who count as *persons* for moral purposes' on the other, those who fall (as proper objects) under the principle of 'respect for persons' in particular, who I have argued need not be 'moral persons' themselves.¹²⁸ Moral persons (human persons) are still special in that it is they who alone are called to act with respect, and also because they are themselves to be respected in a special kind of way, but by no means do they exhaust the domain of moral respectability; human beings are not beneficiaries of the duty that we have not to be cruel to them (and even, I would argue, the duties that we have not to murder them and not to steal their property, which in truth are grounded in practical love and so contain a special reference to persons as intended beneficiaries) because they are capable of morality, for instance.

Now, what I have been doing all along is arguing that non-human animals have a specific kind of moral standing (grounded in our duties to them) which *entails* their personhood for moral purposes. However, I have never argued that animals are 'persons' to begin with *because* they have this moral standing which is grounded in our duties. Again, my claim is that there is a certain moral status X that entails personhood S , but all I mean by this is that X *presupposes* S , not that it constitutes it or grounds it. Animals (i.e. non-human animals) are persons *because they are persons*. Non-human animals are persons because they are subjects, because they have souls or minds, because they are alive in the most basic way in which people (regarded as people instead of as biological organisms) can be alive, because they are personal presences in the world, each a 'world' in their own right. These persons have a special kind of moral significance in our eyes, there is an irreducible

¹²⁸ What I am saying is of course very similar in structure to what my 'adversaries' may be saying at this point, namely: don't conflate 'persons' (in the sense of *psychological subjects*) with those who count as 'persons for moral purposes', which of course have to be 'moral persons' themselves.

moral bond uniting us (moral beings) with them, and so we count them as persons *for moral purposes*. But of course they were persons already, before even being considered in this light: in the world as we know it, the presence of persons precedes the presence of morality and moral relations in general.¹²⁹

You can – looking at my argument as a whole – see how I decided to first build up to the notions of ‘practical love’ and ‘recognition respect’ before emphasising the *personhood* of all animals. I took this to be the correct route to take insofar as the purpose of *moral philosophy* is to explain the *moral relations* that obtain between *persons as such*, i.e. it was my main job to explain why *sentient beings as such* can – and in fact do – participate as objects or recipients in irreducible and universally pervasive (and in these ways *fundamental*) moral relations, which takes substantive moral theorising and cannot be solved by, say, pointing to the etymology of the word ‘animal’, or to sentience as already making a subject or a person (i.e. a ‘someone’ as opposed to a mere ‘something’). But one doesn’t really need to build up in that way in order to arrive at the basic conclusion that *all animals are persons*, and that therefore non-human animals are persons in their own right. Thus, if you think about it, pathological love (in the sense of kindness or benevolence *from sympathy*), which is *not* of itself a moral relation, is nevertheless always and by necessity directed at *persons*. One cannot be kind (whatever the grounds) to a ‘thing’. It is always ‘someone’ that is the object of kindness, whatever the provenance of the kindness. In fact, we can dissociate ascriptions of personhood *even more* by talking about things like fortune and misfortune without connecting them with morality and immorality. A misfortune (calamity, tragedy, etc.) is always *someone’s* misfortune, and the same goes for fortune! And of course: Pleasure is always *someone’s* pleasure! Pain is always *someone’s* pain! ‘Things’ have no knowledge of heaven and hell, only persons do.

We have basic and powerful evidence (philosophical evidence) for all animals deserving to be called *persons*. We have arguments that appear incontestable from the perspective of our reflective common-sense and our everyday parlance (which of course has

¹²⁹ Similarly in the life of a single individual. An individual’s growing from state of ‘bare animality’ into the moral condition of ‘humanity’ is not the same as her growing from a non-person to a person, but from a non-moral person to a moral person.

sometimes been adulterated, although in ways that can easily be spotted and amended). I have given some of those arguments already (e.g. those in the preceding paragraph), and I won't miss the opportunity here to say more: (1) We can only be kind to *someone*, never to a mere 'something'; but it turns out we can be kind to animals (i.e. to non-human animals), therefore animals are persons. (2) Happiness is always someone's happiness; but it turns out that animals can be happy, therefore animals are persons. (3) A person is the kind of entity that can participate in moral relations; but it turns out that animals (i.e. non-human animals), very much like human beings, participate as direct recipients in practical love (a fundamental moral relation), therefore animals are persons. (4) And perhaps the most basic one from the perspective of our common use of language (English being a somewhat complicated case given the frequent - officially enjoined, but in truth completely inadequate as demonstrated in common parlance itself - use of the impersonal pronoun *it* in reference to non-human animals, which in reality is only in reference to 'non-pets' because it is idiomatic to call 'pets' by personal pronouns, and hence the problem that we face: language being commonly used to de-personify those towards whom human beings don't have a personal, pathological, attachment): persons are the kind of entity that are rightly called by *personal pronouns*, but it turns out that I rightly call my cat (my daughter, that is) *she*, and the chickens and cows and pigs and ducks that human beings have killed by the billions *they* (which I do without ever having met them in person), therefore animals are persons. These basic arguments I lay down for the consideration of the reader, and now it is time for a separate (though closely related) point.

All well and good so far. But surely there is a clear *inconsistency* or *irregularity* that has arisen in my story. I began equating 'animality' with the 'capacity for happiness' or 'sensitivity' more generally (i.e. the capacity for happiness + self-love as concomitant agency); but now, and all of a sudden, I have moved in this section to saying that 'animality' is to be equated with 'personhood' as the very same philosophical category, which seems clearly to be changing the rules of the game. If you recall, all my arguments up until the last section (and even the ones in this section) have relied on the thought that there is more to human beings *as persons* than just 'animality' or (what I have taken so far to be the same thing) 'sensitivity', and this alone entails that 'animality' (i.e. if equated with 'sensitivity') cannot equal 'personhood'. Moral personality is an irreducible aspect of persons *as persons*,

and this alone is enough to show that the two proposed uses of the word 'animal' are mutually inconsistent. What is it, then? Let me explore the two options that we have. Ultimately, there is a third which I will try to motivate.

The first option is to *undo* the association between 'animality' and 'sensibility' (happiness + self-love) and insist on 'animality' and 'personhood' being one and the same philosophical category. This seems right, insofar as the basic etymological meaning of 'animality' is 'having a soul' or 'having a mind' (i.e. souledness or mindedness), and it is eminently plausible to equate 'persons' with 'those who have a soul' or 'those whose have a mind'. In fact, this might well be the most basic meaning of *person* as a concept. Then we could of course think of 'persons' or (what is the same) 'animals' as having one or more types of souls (one or more types of mindedness if you will): persons / animals may be both 'sentient' *and* 'rational' OR they may be only one of those two things, thinking of those who only instantiate 'sentience' (who we have encountered) and not so much of those who allegedly instantiate 'reason alone' (whom we have not). If the latter existed (i.e. assuming that they can even exist metaphysically speaking, which is a separate question entirely), we could rightly call them 'purely rational persons' or 'purely rational animals' or 'purely rational subjects' or 'carriers of a purely rational soul'. Again, there are those who only have sensitive souls and concomitant agency (non-human animals) and those who have both rational and sensitive souls (human animals), where both are assumed to be sentient and to count as 'animals' and therefore also as 'persons' merely on that account already: any of the two types of 'souls' - and of course also the two together - is enough to make an 'animal' or a 'person'.

The other option starts by insisting on 'animality' standing for 'sensibility' in particular (happiness + self-love as concomitant agency), which initially presents itself as problematic from the perspective of equating 'animality' and 'personhood' (as I came to do only very late in the argument) because there seems to be more to 'personhood as such' than just sensibility plus concomitant agency. Again: surely there are persons with more to them *as persons* than just 'sensibility' (capacity for happiness + self-love) and potential modifications of it; but since I am now committing myself to 'animality' equating 'personhood', then I seem to be suggesting that the addition of a 'rational soul' would of itself contribute nothing to the 'personhood' of those particular persons, which just seems

wrong. We do talk about ‘moral personality’ and take it to be an irreducible aspect of the person (the human person) *as such*, certainly irreducible to what I have so far deemed ‘the animal condition’ (the capacity for happiness + self-love). Hence, if one wants to insist on preserving the identity between ‘animality’ and ‘sensitivity’, one seems obliged to undo the claim (my late claim) equating ‘animality’ with ‘personhood’ to avoid unduly reducing ‘personhood’ to ‘sensitivity’. Or perhaps it is possible to find a third way.

I have to be honest and admit that this analysis is not meant to be final, as I am still undecided about whether to *undo* the claim that ‘animality’ equals ‘sensitivity’ (the capacity for happiness + concomitant agency), which of course would have to be done elsewhere. What I am certainly not undoing is the claim (which I came to make late in the argument) that ‘animal’ equals ‘person’ equals ‘subject’ as the same philosophical category. It is one of the scandals of the history of philosophy, I think, that it brought us to a place where ‘animality’ (the original meaning of which is ‘having a soul’) is commonly associated with ‘corporality’ or ‘belonging in the domain of nature’ and so by implication not in the that of ‘Freedom’ or ‘Reason’ *as the only one that is inhabited by ‘persons’*.¹³⁰ Freedom and Reason,

¹³⁰ I would like to quote Kant in relation to this point (MM 6:223): “A *thing* is that to which nothing can be imputed. Any object of free choice which itself lacks freedom is therefore called a thing (*res corporalis*)”. Kant is here quite explicitly placing non-human animals (who lack freedom just as much as inanimate entities lack freedom, and to whom nothing can therefore be imputed) in the domain of ‘things’ (*res corporalis*), which contrasts with the domain freedom as the only one inhabited by persons (MM 6:223): “A person is a subject whose actions can be imputed to him. [...] From this it follows that a person is subject to no other laws than those he gives to himself (either alone or at least along with others).” To be clear, Kant doesn’t think that non-human animals are devoid of sensation, experiential subjectivity or even (non-moral) volition, being quite emphatic in ascribing them all of those things. He doesn’t equate ‘animal’ with ‘corporeal’ either. But as we can see clearly in the lines just quoted (as well as in many other than have been quoted), he does speak in ways that entail the lack of personhood (and membership in the merely corporeal world: *res corporalis*) of non-human animals in virtue of their lack of freedom, which is problematic in light of the very etymology of the word *animal* (which designates the concept of *souledness* or *mindedness* and therefore contrasts with the concept of the merely corporeal if anything does). For Kant, then, there is more to the ontological map than just freedom and non-freedom, reason and non-reason: there is also sentience (which for him is very close to animality, as it has been for me up to this point) and everything that comes with it; however, if we are drawing an ontological map for the sake of situating *persons*, we need only draw the limits between freedom and non-freedom, between reason and non-reason, and have no need whatsoever to talk about anything else. Now, it

I agree, we can reserve for human beings, but *personhood* (an entity's *being a person*, i.e. an entity's *being someone* as opposed to a mere *something*) belongs to animals in general, and this *by definition* and therefore obviously and indubitably. The task is in our hands, therefore, to rehabilitate 'animality' and 'animal' as the central philosophical categories that they are.

7. Utilitarianism & the Vindication of Kantian Ethics

Even though I have been defending the dignity of all animals / sentient beings (as a good Utilitarian would), I have been doing so without ever positing happiness and its bearers so regarded as the ground or locus of moral realities (as a good Kantian would). The thought has remained intact that the concepts of moral duty and moral goodness are answered only by intrinsic features and modes of our own *humanity*, of our own *moral agency*, without any reference to the felicitic properties of our conduct. This is not bad news for animals, of course. The existence of humanity itself is meant to be good news for animals in general, since it is intrinsic to the task of humanity to exercise practical love for all animals: it is beyond us (i.e. beyond our morality) that they actually be happy, but still we are here to protect them as our kindred and do our best effort for their happiness. That is of course as much as we can promise them. But I have said enough about the moral standing of animals already, and need to shift focus towards *humanity* as such, which is what Utilitarianism (as the regular ethical theory that I take it to be) is primarily all about.

7.1 Gathering Historical Pieces

I don't think I need to repeat the slogans with which the theory known as 'Utilitarianism' has often been associated. I already talked about them in **Part I**, where I showed that the

is not only Kant who is responsible for this scandal, as many other people (many other classics, e.g. Descartes and Locke, and many non-Kantians, e.g. most Christians, including virtually all 'conservative Christians') have painted the exact same picture.

basic formulations of the view following those slogans are fundamentally flawed, in particular the ‘Consequentialism’ + ‘Welfarism’ formula that has become dominant in philosophy. My purpose here is to restate Utilitarianism.

To kick things off, let consider the following line by John Stuart Mill, extracted from his *Utilitarianism*: “the morality of the action depends entirely upon the intention – that is, upon what the agent *wills to do*” (Mill 1861/2003, p.196).¹³¹ Read it again if you need to. Is it so patently clear – I ask – that Utilitarianism is intended as a form of ‘Consequentialism’ that as such stands in opposition to theories making the morality of actions depend on properties of the intentions or volitions of human beings? If so, Mill apparently didn’t know about it. And if you think – because that is the rumour – that Mill was not in fact a ‘Utilitarian’ but instead Judas among Utilitarians precisely because he says things like that, now read the entire line (Mill 1861/2003, p.196):

There is no point which utilitarian thinkers (and Bentham pre-eminently) have taken more pains to illustrate than this. The morality of the action depends entirely upon the intention – that is, upon what the agent *wills to do*. But the motive, that is, the feeling which makes him will so to do, when it makes no difference in the act, makes none in the morality: though it makes a great difference in our moral estimation of the agent.

Utilitarianism, Mill tells us, makes the ‘morality of actions’ depend on properties of human *intentions*, and therefore not on their felicitic consequences. This is supposed to be true for both Bentham and Mill, taking it on Mill’s word not to make things unnecessarily complicated. Now, we are yet to see what this claim by Mill (to keep focusing on him) really amounts to, but there is another element in the picture already: the distinction – which Mill is positing in the line just quoted – between the ‘morality of the act’ and the ‘morality of the agent’, which is what I am naming the object of “our moral estimation of the agent” (Mill 1861/2003, p.196). Now, the reason Mill is positing this distinction is because he is trying to convey the point that there are two ‘moral judgements’ that come (or may come) apart depending on the features of the situation that we focus on: when we are ‘judging actions’, he says, we focus on *intentions*, but when we judge persons (a.k.a. human beings)

¹³¹ This is Mill’s own emphasis.

we instead focus on their *motives*. Now, here I don't care about possible ramifications Mill's distinction or even about its merits, as the point is simply to note that *both* of the moral qualities (both of the 'moralities') that Mill thinks we assess in human activity – both of them, I repeat – are said to be entirely fixed by properties of *human agents*, namely their *intentions* and their *motives* respectively.

Now, even though I celebrate that Mill (to keep speaking of him) is making the 'the morality of actions' depend entirely the *intentions* of human beings (to start with this particular point), one is initially disappointed at realising how *thin* a conception of human intention Mill seems to be working with. In his writing on Bentham he tells us that the "morality of an action depends on its *foreseeable consequences*; its beauty, and its loveableness, or the reverse, depend on the qualities which it is evidence of" (Mill 1861/2003, p.84).¹³² What Mill is saying here seems again to correspond to the distinction between the 'morality of the act' (which depends on 'intentions') and the 'morality of the agent' (which depends on 'motives'), only that now he is telling us what he really takes 'intentions' to be: foreseeable consequences? So, is Mill now conceding that it's all about consequences, only not in actuality but in some other form? I would say that he is, but he is not thereby retracting the basic point about human intention. After all, talk of 'foreseeable consequences' or 'expected consequences' or 'intended consequences' is always to be construed as talk of *human intentions*: very thin intentions to be sure, intentions that so far seem to be grounded in the air, *mere pursuits of outward outcomes* if you will, but intentions nonetheless. But then of course there are also *motives* in the picture, as Mill says, and *these* could well ground the initially very thin intentions he was talking about.

Now, I would have been neat if at this point Mill had said: *from duty, for everyone's happiness*, or just *from duty* in case he wanted to simplify (or just *for everyone's happiness*, as one could also rightly simplify). It would have been neat of him to point to some determinate principle or maxim or motive the influence of which, in human action, would entirely fix the *morality* (positive moral worth) of that action. But no! Instead, he tells us that 'motives' are immaterial to the 'morality of actions', and that it's only *intentions* (i.e. 'bare intentions', intentions thought of in isolation from their grounds and even from their

¹³² My emphasis.

actual consequences, although still in direct relation to their ‘intended consequences’, which of course doesn’t detract from their status as *intentions*) that count. Again, it is the initially very thin intentions that ‘count morally’ in Mill’s view: they are grounded in motives, to be sure, but it’s not this connection that makes them ‘moral’. Or wait a second: is Mill really talking about *morality* in the strict sense of the term when he makes these claims? Or is he perhaps talking about something else? Let me quote another line that I find instructive (Mill 1861/2003, p.195):

He who saves a fellow creature from drowning does what is *morally right*, whether his motive be *duty*, or the hope of being paid for his trouble.¹³³

Mill calls it the ‘morality of actions’, but of course it is ‘moral rightness’ that he is talking about, which he distinguishes from ‘the morality of the agent’. Is Mill just giving us Kant’s ‘legality vs morality’ distinction, only in other words? If he isn’t, or (in any case) if it is really his view that helping a fellow ‘for the sake of getting a reward’ counts as a *genuinely moral* action, then of course he is in big trouble. It would be a very silly, almost unbelievable, mistake to give us a moral theory calling actions ‘moral’ (genuinely moral) that are not motivated or driven by any kind of moral consideration. More than a silly mistake, it would look very much like a cheap attempt at denying the concept of morality. If, however, we decide to map Mill’s ‘moral rightness’ onto Kant’s ‘legality’ or ‘rightness’, we get a completely different story: both philosophers can agree in saying that a human being helping a fellow from ‘purely self-interested motives’ can in principle be called ‘right’. If you look at the person acting and all looks well *from the outside* (at least as pertaining to her *apparent intentions*, for ‘actual consequences’ are completely immaterial at this level, and this even for Bentham and Mill), i.e. if we look at the situation and the person whose conduct we are assessing seems to be ‘doing the right thing’, then we can all agree to call it ‘morally right’ or ‘legal’ or ‘in external conformity with moral commands’. Certainly we have no grounds for calling it ‘wrong’. And this is just the first potential point of agreement, of course, since we are yet to say something about what Mill calls the ‘morality of agents’,

¹³³ Both emphases are mine.

which he tells us depends entirely on motives ('duty' being the one that he mentions in the context of the example we are discussing) and which we could well be mapped onto Kant's concept of *morality*. Accordingly, we can read Mill as saying that the action of saving a fellow is *morally good* (genuinely moral) only in case it is committed 'from duty' (or in any case from a distinctively moral motive or principle or maxim). And voila! Have we put an end to the biggest battle in moral philosophy?¹³⁴

As far I am concerned, Utilitarians are perfectly free (i.e. we don't have to abandon our view) to depict morality as irreducibly a matter of attitude on the part of the moral agent herself, of this agent voluntarily aligning her conduct with the internal standards (commands) of her own humanity (autonomy - human dignity - Kantian ethics). And the same with Kantians, who we are perfectly free (i.e. we don't have to give up our Kantian label) to state the above and then add that morality is also at the same time (both essentially and irreducibly) a matter of *self-givingness*, i.e. of the agent voluntarily turning herself into a means for the happiness of subjects (Utilitarianism). If obeying reason within can be shown to involve reaching out to sentient beings (acting for their happiness) as a fundamental moral act of the human being, then the conclusion just follows that there is no real fundamental conflict between Kantian ethics and Utilitarianism.

To be clear, practical love and the self-givingness (the 'acting for the happiness of subjects') that it essentially involves are already in place, and I don't need to be adding any arguments to that effect. What I am setting myself to do at this point is elucidate the nature of what I take to be the *distinctively moral attitude* (the attitude whose instantiation alone constitutes *moral goodness*) in its full moral relationality. So far we have the slogan: *from duty, for everyone's happiness*, but a more elaborate argument can be made putting various ideas together in something like a marriage between the following three views (taken as theoretical idealisations or specific philosophical theses, not necessarily as 'traditions in philosophy'): First-Personal Kantianism (Kantianism emphasising *self-legislation*),

¹³⁴ At this point it is being assumed that Utilitarians need not (and in fact should not) be committed to 'Welfarism' as their *moral axiology*. The story that I have been telling in this **Part II** (and that I will continue to tell in the next subsection, now in more technical fashion) is a story about how to understand Utilitarianism as a theory about the nature of *moral goodness* in the strict Kantian (realisation of the human being as a moral being) sense of the term.

Contractualism (Kantianism emphasising *co-legislation* and related human-to-human moral relations, the most general one *in my view* being that of *co-guardianship in the moral community*), and Utilitarianism (practical love for all sentient beings). This corresponds to my technical formulation of Utilitarianism, a label that I am claiming for my moral philosophy *as a whole*.

7.2 On Utilitarianism (and Kantian Ethics): The Technical Formulation

In virtue of the relationally-specific duty of practical love which permeates the entirety of morality, the motive of duty acquires the sense of *self-givingness*, of doing things *for the benefit of subjects*. ‘Practical love’ is love (benevolence, kindness) *from duty*, but it is also always at the same time ‘self-giving love’, and thus the possibility to define *Utilitarianism* (the theory making morality a matter of practical love) as the view according to which *everything* that is done truthfully in the name of ‘morality’ or ‘from duty’ is done always at the same time *for everyone’s happiness*. If an action doesn’t come from the exercise of a maxim that is adopted *for everyone’s happiness* (i.e. from practical love), the action consequently lacks *moral worth*. Now, I have made it very clear that none of this is to deny the ‘more basic’ Kantian slogan for the morality of actions: only actions that are done *from duty* – and only because they have that feature – can rightly be said to be *morally good*. Here what we are doing is simply further describing what *acting from duty* amounts to.

I take all the work that has been done to speak for the recognition of the double relationality expressed in the slogan: *from duty, for everyone’s happiness*, but there is still more to be said to that effect (i.e. to justify the second part of the slogan in particular as the one that is telling us something ‘new’ about the nature of the motive of duty). Let me then try to give you something closer to a proper argument for *morality as a whole* (and therefore the motive of duty as the distinctively moral motive) essentially involving an element of *self-givingness* that justifies our basic slogan and the use of label ‘Utilitarianism’.

If Utilitarianism presents itself as the philosophy of ‘practical love’, a philosophy devised to emphasise the fact that the *whole* of morality is done ‘for the happiness of subjects’, does that mean that such a philosophy is bound to undo or reduce all the other moral relations we have been talking about? Didn’t I affirm the notion of ‘respect for

human beings *as such*? And is there not an intra-personal dimension of morality (self-legislation) that has a good claim to being counted as the most fundamental one? And if I don't mean to undo what I have said, and if I am not here saying that we should reduce everything to 'practical love', and if I am myself portraying practical love as only a *part* or an *aspect* of morality alongside other fundamental parts or aspects, wouldn't it be more plausible to say that only this one part of morality, and not so the whole, is done *for the happiness of subjects*? And wouldn't it be more plausible – for the very same reason – to say that 'Utilitarianism' is trying to account for a particular aspect (as opposed to the whole) of morality?

Let me introduce more nuance into my argument. Initially, 'practical love' is meant to be understood as a fundamental, if only *relationally-specific*, moral relation. By calling this moral relation 'relationally-specific' what I mean to indicate is that it has a *specific* direct object, which in this case is *sentient beings in general as intended beneficiaries*. But this is not all that there is to say about the nature of this moral relation, of course: 'practical love' is also to be understood as a moral relation that is *universally pervasive* of morality, a relation that features as *an essential aspect of every moral duty that we have*. This universal pervasiveness, I argue, is enough to ground the claim that *the whole of morality* (in the sense of *every single moral duty that we have*) essentially involves an act of *self-givingness* on the part of the human being: in virtue of 'practical love' as ground, every moral duty that we have is a call to act *for the happiness of subjects*. And this in turn is enough – or so I argue – for me to claim the *Utilitarian* label for my proposed moral philosophy *as a whole*: morality is, taken in one piece and therefore *every bit of it*, a matter of practical love for all sentient beings.

Think about it. A similar story *has to be told* for other essential aspects of morality that people also like to emphasise. Why is the whole morality (i.e. every single moral duty that we have) a matter of obedience to reason within (Kantian autonomy – self-legislation)? Is it because obedience to reason within is the only fundamental moral relation, or is it because this relation is universally pervasive of morality? I think it has to be the latter: it has to be because *every moral duty that we have is self-imposed* (and not because the self-imposition of moral duties is the only fundamental moral relation) that the whole of morality (the whole set of our moral duties and so every step that we take in the name of morality) is a matter of obedience to reason within. The universal pervasiveness of obedience to reason within

(the fact that *every duty is self-imposed*) makes it the case that morality is, taken in one piece, a matter of obedience to reason within, of voluntarily aligning one's conduct with the internal standards of one's own reason as the only moral standards one is ethically obliged to live up to. And making the picture complete, the same goes for the 'Contractualist' aspect of morality: the idea of morality being a matter of 'human collaboration' (moral co-operation involving co-legislation and co-guardianship of the moral community of which all sentient beings are intended beneficiaries). The universal pervasiveness of this 'third' aspect of morality, it's being essential (if merely as an aspect) to every moral duty that we have, makes it the case that morality is – taken in one piece – a matter of human collaboration: there is not a single duty that we have that is not always at the same time a matter of human collaboration, of 'doing one's bit' in the moral community *vis-à-vis* other human beings.¹³⁵

This analysis is meant to speak in favour of claiming the label 'Utilitarian' for my moral philosophy *as a whole*, and I really think that is what it gives me. But then again this point also speaks for other labels as well, thus bringing me back to the very place I was trying to escape. The fact that Utilitarianism has to stand alongside other labels sort of speaks against Utilitarianism as 'the correct moral theory', doesn't it? Well, it doesn't. It speaks in favour of our *not* having to choose one label above another, which in a way is the entire point of this philosophical work. Let me however try to elaborate more around the same ideas to see if further philosophical progress can be made allowing us to put all the pieces of this 'puzzle' together.

¹³⁵ This is *not* the idea that one is to divide the pie of responsibilities (as it were) fairly into slices among all human beings and only do 'only one's bit': that if – for instance – there were two children drowning and another capable human being around who could help out but (for no apparent good reason) decides not to, I would need only save one child and let the other drown. In a case like this, saving both children (and *perhaps* – depending on details of the case – doing something on top of that to address the human being who didn't help out) constitutes 'doing one's bit' (i.e. one's job) *vis-à-vis* human beings in general. Is there unfairness in the situation in which one acts this way *in part because one acts this way*? Yes, but in saving the two children one is not oneself being unfair, one is in no way putting benevolence above fairness. It's the other human being who (presumably) is being malevolent (to put it like this) and *unfair*. It is him who is responsible for the unfairness (the strictly moral unfairness) of the situation.

I won't deny that there is a sense (a crucially important sense) in which the whole of morality is owed *only* to reason (the moral law) within: the *obedience* that morality essentially involves is *as such* owed (ethically speaking) to reason within and to reason within alone – we don't owe obedience (*at least not directly*) to our fellow humans as such or to sentient beings in general. Likewise, it has to be admitted that there is a sense in which the whole of morality is owed to sentient beings so regarded and to sentient beings alone: the *self-givingness* that morality essentially involves is as such owed (ethically speaking) to sentient beings in general as the intended beneficiaries of the moral enterprise – we don't owe self-givingness (*at least not directly*) to reason within or to human beings so regarded in general. And again likewise, and to bring completeness to the picture, *human collaboration* is as such owed (ethically speaking) to other human beings as moral co-workers in the moral community, and to them alone – we don't owe this collaboration (*at least not directly*) to reason within or to sentient beings in general.

Now, the key to getting the point that I am trying to make lies in understanding what I mean in every case by the 'at least not directly' clause. More precisely, it lies in understanding why the point should not be made using that clause, but in a slightly different way. A lot of theorising in this area makes use of the expression 'owing things directly' and 'indirectly', but I argue that we should avoid it on the grounds that it is not very illuminating and even potentially misleading, i.e. supposing that it even makes sense to say things like: *we owe obedience to our fellow humans and sentient beings in general, but only 'indirectly'*. It is statements of this kind, statements using the expressions 'owing things directly' and 'indirectly', that tend to reinforce the reductionist thoughts of philosophers who are in the business of defending certain labels as against other labels. The 'reductive first-personal Kantian', for instance, will probably take the chance to once again portray people like me (who argue for duties to animals) as confused because clearly morality is 'most essentially' a matter of *obedience*, and obedience can be 'owed outwardly' (e.g. it can be directed to human beings or sentient beings) only 'indirectly' *and therefore not really* (which would be odd for a 'Kantian simpliciter' to say anyway given that it would make duties to other human beings 'indirect duties', which most likely he doesn't really want to say; although he will probably find a way to include human beings in general that doesn't apply to other animals, who, for instance, are not bound by obedience themselves). You

can also imagine the ‘Reductive Contractualist’ taking advantage in a similar way, and the very same for the reductive Utilitarian (Agapist) out there: no matter how adequately nuanced one’s claims are otherwise, if one makes them in terms of owing things ‘directly’ or ‘indirectly’ one immediately gives way for philosophers to bring their own idiosyncrasies to the fore in a contest (a very dangerous contest) for who is the one ‘with the right answer’. Let me then avoid using the terms ‘directly’ and ‘indirectly’ and rephrase my argument in more adequate terms. Let me use the language of ‘owing duties to *X*’ (*X* being the relevant party at the ‘receiving end’ of whatever fundamental moral relation we are talking about) *in a relationally-specific way*. If we use this terminology instead (which I have already been employing in reference to ‘relationally-specific attitudes’: fundamental moral attitudes with their own specific direct object), all confusion begins to disappear and ideas begin to really come together.

Relationally-specifically, the kind of *obedience* that morality essentially involves is owed only to the Moral Law (reason) within, meaning that it is obedience *to* reason within and to nothing else (no ‘indirectly’ here). In the relationally-specific sense that is implicated here, we don’t live for anyone other than ourselves: we are not here to ‘please’ anyone, only to do as our own reason says. Relationally-specifically, the *self-givingness* that morality essentially involves is owed only to sentient beings as such and in general, meaning that it is self-givingness *directed towards* sentient beings. In the relationally-specific sense that is implicated here, we very much live *for everyone*, which is not to deny the above because giving oneself to sentient beings in practical love is part and parcel of doing as reason says. Again: *from duty, for everyone’s happiness*. And something similar goes for *moral collaboration*, of course, which (speaking in a relationally-specific way) we owe to our fellow human beings as co-workers in the moral community and to them alone under that condition, not to reason within or to sentient beings in general (again, no ‘indirectly’ here, as the point here is simply to note that moral collaboration is collaboration *with* other human beings).

Now, this all seems again to be speaking in favour of keeping Utilitarianism (practical love and self-givingness) restricted as a particular aspect of morality, instead of ‘morality as a whole’. In some sense, this is true, but then again I see it as a good thing because from the very beginning the whole point has been to depict Utilitarianism as consistent with other labels in philosophy.

But there is still one last effort that I can do to make Utilitarianism (or in any case Agapism) even more central than I have managed to make it thus far. This last effort may involve problematising some of the claims that I myself have made, in particular the claim that practical love is just one aspect of the fundamental moral principle as opposed to the principle itself, but that is exactly what this last effort of mine consists in. If the principle of ‘respect for persons in general’ can as such be called ‘the principle of practical love’, then Utilitarianism (Agapism) must have some genuine claim as a label for our moral philosophy *as a whole* (or perhaps the main label as among those that apply to the moral theory as a whole).

This is how I see things going. Even though we don’t owe ‘obedience to reason within’ to sentient beings in general *in a relationally-specific way* (for they are obviously not the direct object of this attitude), we can still talk about owing them (as intended beneficiaries) *all acts* that constitute obedience to reason within. This is of course the very same point I have been making all along: all moral acts stand in relation to reason within as moral authority (as the only source of their obligatoriness), but this doesn’t prevent those same acts from standing in relation to sentient beings in general *as intended beneficiaries*. But of course this *double relationality* is assumed by the very *definition* of ‘practical love’ that I have been working with all along. If it doesn’t come from morally legislative reason but from some other place, if it doesn’t belong in the idea of living up to the internal standards of humanity, it just doesn’t count as practical love. Such is exactly what distinguishes ‘practical’ (rationally-based) from ‘pathological’ (sensitivity-based) love, remember. And so I ask: how is it that adding other human beings in the picture as co-legislators and co-workers in the moral community (the Contractualist aspect of morality) can possibly ruin the status of the basic moral principle as a ‘principle of practical love’ if we have agreed that reference to the intra-personal dimension of morality cannot do that? It can’t, of course, and so I say: if it doesn’t put the moral subject in relation with other moral subjects as co-workers or co-guardians in the moral community, if it relates to human beings as ‘mere intended beneficiaries’ or ‘passive citizens’, i.e. if it treats human beings as babies, it ain’t practical love. If it doesn’t involve truthfulness of testimony (especially human or moral testimony, i.e. calling things by their name) *vis-à-vis* human beings in general (as well as sincerity in promises and other things that may rightly belong in this dimension of

morality), it just ain't practical love. But if – given this basic clarification – morality can in fact *as a whole* be seen as a matter of practical love, i.e. if this 'practical love' can rightly be deemed uniquely necessary and sufficient for *moral goodness* to obtain (because it embodies direct obedience to the moral law and respect for persons in general, both as capable of morality and as capable of happiness), then we really have to entertain the possibility that Utilitarianism (Agapism) – the theory depicting morality in this way – is actually true. But of course this is also my proposed reconstruction of Kantian philosophy, and I take it to follow as a conclusion of all the work done that there is no real conflict between Kantian ethics and Utilitarianism construed as an *Agapist* ethical theory.¹³⁶

¹³⁶ At the very least, it has been shown *how* (and by implication also *that*) it is possible to make meaningful and systematic use of both of these labels in the context of accounting for the basic facts about the nature of morality (including the basic map that we need to draw of the fundamental moral relations that obtain between persons in the moral community).

Part III (Conclusion)

Let me bring this work to an end by offering further reflection concerning the relation between Kantian ethics and Utilitarianism, only now more so from the perspective of what they share (or can be seen as sharing) as *historical movements in moral philosophy*. Real differences between them notwithstanding, there is one common feature in relation to the genesis of these movements that we do well to acknowledge, namely that were developed under the maxim of: (i) making good of ancient – especially Hellenistic, i.e. Stoic and Epicurean – philosophy, and (ii) secularising – i.e. divorcing from theological *foundations* – the fundamental ethical insights of ‘Christian philosophy’. Let me begin by quoting Mill’s *Utilitarianism* again (Mill, 1861/2003, p.187):

I do not, indeed, consider the Epicureans to have been by any means faultless in drawing out their scheme of consequences from the utilitarian principle. To do this in any sufficient manner, many Stoic, as well as Christian elements require to be included.

Before developing my claim, I need to note that Mill writes these lines in the context of his famous (or should I say infamous) argument to the effect that ‘Utilitarianism’ (as allegedly a ‘hedonistic’ theory that because of that relates very closely to Epicurean philosophy, which I argue is a mistake if taken to mean there is no ‘moral goodness’ irreducible to the felicific or the hedonic – which I *doubt* is what Mill really means to say anyway, but in any case) is not a ‘philosophy for swine’ because it acknowledges – as he takes ancient epicureans to have acknowledged – that human beings have access to ‘higher pleasures’ (pleasures of the intellect and social interaction) which far surpass in ‘value’ – as components of happiness, i.e. in *felicific* value – that of the lower pleasures also available to non-human animals, and that therefore human happiness is almost incomparably better (again, in the strict *felicific* sense) than ‘merely animal’ happiness, and that therefore our human ‘sense of dignity’ comes out vindicated in ‘Utilitarianism’. This is the weakest argument in the entire book, I think. Not only is Mill’s claim false that ‘higher pleasures’ are as such more valuable (felicifically speaking) than ‘lower pleasures’ (as not only Bentham but also Kant rightly

pointed out),¹³⁷ he is also completely mis-locating the ground of our special ‘sense of human dignity’. As Kant emphasised with the Stoics and Christian philosophy in mind, what makes human beings special is our moral vocation as such, the fact that we are called for morality, which we are in such a way that self-love (i.e. our natural concern for our own happiness, or our own ‘pleasure’ as Stoics who think ‘happiness’ lies elsewhere would always have to call it) has to take second place in our practical life. In a certain sense, morality is very much all about subordinating the call of self-love to the call of moral duty, to constraint the pursuit of our own happiness by the moral pursuit (which – as Kant himself makes it very clear – essentially involves the pursuit of the happiness of subjects in general). This is of course the essence of the Kantian story, which Mill obscures by trying to place the ground of human dignity elsewhere, in a place that in truth we are the equals of other animals (as in a certain sense even Kant acknowledges).

But are we supposed to conclude from this analysis (of the argument in the context of which Mill invokes ‘Epicurean’ and ‘Stoic’ and ‘Christian’ philosophy as having to come together in making sense of ‘Utilitarianism’) that therefore Mill – or Utilitarianism itself – lacks the resources to be able to follow the Kantian story? If I understand Mill’s intention here, he is invoking the two schools more associated with virtue than they are with pleasure (the ‘Christian’ and ‘Stoic’ schools) because he wants to point out that ‘Utilitarianism’ is not out to portray human beings as having no more mental powers than pigs do; but of course that is a point that *all moral philosophy* has to assume in the idea that human beings – and not so others animals, i.e. to the extent that they are really *non-human* – are moral beings capable of bearing duties and being morally good, etc. If Utilitarianism is to be presented in a fair-minded way as an ethical theory, there is absolutely no need to appeal to felicific distinctions to defend *human dignity* in the (also fair-minded) Kantian sense. Now, I am of course focusing on the negatives of Mill’s efforts (seizing the opportunity to make a general point about ‘dignity’ that once again captures the essence of this work); but actually this is not the only place where reference to ‘Stoic’ and ‘Christian’ philosophy comes up in Mill’s *Utilitarianism*, and I think we do well to examine those other references because they speak to something good and deep.

¹³⁷ I will come back to this point.

We can safely assume that the ‘Christian’ component (or perhaps the most important ‘Christian’ component) in Mill’s narrative has to do with the fundamental moral principle itself, for he quite explicitly tells us to make this connection: “In the golden rule of Jesus of Nazareth, we read *the complete spirit* of the ethics of utility. To do as one would be done by, and to love one’s neighbour as oneself, constitute the ideal perfection of utilitarian morality.”¹³⁸ (Mill, 1861/2003, p.194). This is perfectly consistent with the story that I have told in this work, construing Utilitarianism as an *agapist* doctrine that therefore belongs in the domain of ‘human dignity’ in the Kantian sense, thus making it the case that Utilitarians have at their disposal the same answer that everyone else (e.g. ‘Stoics’ and ‘Kantians’ and ‘Christians’) gives when asked if human beings are capable of something ‘higher’ than pleasure: yes, they are capable of self-legislation, co-legislation, co-loving of creation, speaking for the ‘voiceless’, impartial self-givingness for the happiness of subjects, etc., i.e. human beings are capable of *morality*. More recently, Peter Singer made a similar association (Singer, 2017, p.31):

All of these philosophers¹³⁹ can be seen as presenting utilitarianism as the best understanding and application of the insight that underlies the Golden Rule. Nor is it an accident, we believe, that something akin to the Golden Rule lies at the core of the ethics of many different cultures and civilizations, from the Jewish and Christian traditions to those of India and China. That utilitarianism can plausibly be seen as an implication of the same insight is a further argument in its favour.

Let me stop to consider this claim by Singer, which I find instructive. The fact that in the last line Singer speaks of Utilitarianism as an *application* of the moral insight embodied in the ‘Golden Rule’ suggests that he is talking about ‘Utilitarianism’ as something – a

¹³⁸ My emphasis.

¹³⁹ Singer mentions Bentham and Mill in relation to their “everybody is to count for one and nobody for more than one” slogan; Sidgwick and the “requirement that we regard the good of any one individual as equivalent to the good of any other”; Harsanyi’s and his idea of the ‘veil of ignorance’ which “forces us to be impartial between all members of the group for which we are choosing”; Smart and his appeal to the sentiment of ‘generalized benevolence’; and lastly Hare and his “analysis of moral language as requiring us to put ourselves in the position of all of those affected by our actions”. (Singer, 2017, p. 31)

procedure such as the ‘felicific calculus’ perhaps – that is called for *given* the aforesaid insight as opposed to something (a moral theory?) giving us the insight *itself*, which doesn’t exactly speak in favour of my point. But that is only the last line in the passage, and right before that Singer also speaks of Utilitarianism as providing an “understanding” of the Golden Rule (Singer, 2017, p.31), which does speak in favour of my point because a theory putatively giving us an understanding of a certain insight (and portraying that insight as the fundamental moral insight available to humanity) is a theory that claims that insight for itself as its very essence, which is exactly what Mill does in relation to ‘the Golden Rule of Jesus of Nazareth’ as one that captures the “complete spirit” of Utilitarianism (Mill, 1861/2003, p.194).

The name ‘Utilitarianism’ may be a new thing, but the ones who first identified themselves by it (and even those who still identify) didn’t see the referent as a new thing at all. If anything, they saw it as one of the ancient belongings of the ‘human species’. Utilitarianism is meant to be a philosophy based on human reason vindicating ‘practical love’ as the fundamental duty of the human being. Insofar as Kant can be regarded in the same (or similar) light, this is an important point of encounter between the two projects in moral philosophy. Now, let me admit that this particular point (the fact that both are attempts to secularise the fundamental ethical insights of ‘Christian philosophy’, and especially – for my purposes – the notion of ‘agape’) is to a large extent philosophical cliché. After all, the only thing I have pointed to here (the ‘Golden Rule’ as registered in a particular expression of it) is something that speaks to everyone but which at the same time admits of many possible interpretations, and nothing can be assumed on this basis about how much the two projects in moral philosophy actually have in common. Now, you can of course fill in this gap with the content of this work, or you can try to fill it in yourself, of course.

I am not trying to diminish the importance of the fact that both projects in moral philosophy try to associate themselves with Christian philosophy. But I don’t think this is the most *theoretically interesting* association that they have in common. If I am allowed, I would like to give this particular honour to Hellenistic philosophy. We already saw Mill invoking both ‘Stoic’ and ‘Epicurean’ philosophy as necessary to understand ‘Utilitarian’ philosophy. Now, his own way of arbitrating between the two ancient schools and putting

them together is problematic in ways that I have explained already: for instance, I don't think Mill is the best representative of 'hedonism about well-being', nor do I think he manages to do justice to human dignity in something like the Kantian sense. But the fact is that Mill partially redeems himself with his reference to the 'Golden Rule', and there is another reference to the Stoics (and in fact to Kant himself under the name 'the Transcendentalist') that I think we need to consider to get the full picture of what he has in mind for Utilitarianism (Mill, 1861/2003, p.194):

Meanwhile, let utilitarians never cease to claim the morality of self-devotion as a possession which belongs by as good a right to them, as either to the Stoic or to the Transcendentalist. The utilitarian morality does recognize in human beings the power of sacrificing their own greatest good for the good of others. It only refuses to admit that the sacrifice is itself a good. A sacrifice which does not increase, or tend to increase, the sum total of happiness, it considers as wasted. The only self-renunciation which it applauds, is devotion to the happiness, or to some of the means of happiness, of others; either of mankind collectively, or of individuals within the limits imposed by the collective interests of mankind.

Utilitarians, Mill claims, are committed to sacrificing their 'greatest good' if morality so requires. You can take this as the resolution to put one's life in line in some 'extreme situation', but I suggest we read it in the more general (Kantian) sense of morality not being grounded in self-love and indeed as involving the *subordination* of self-love to moral duty, death and misfortune come as they may. In fact, I suspect that such is what Mill actually had in mind here. Initially, however, the story that Mill gives us may seem to contradict Kant's at some level: while Kant tells us that self-love is always to be subordinated to moral duty *out of reverence for the moral law* (and in this sense 'directly for the sake of the law'), Mill tells us that 'moral sacrifice' is a *waste* if it "does not increase, or tend to increase, the sum total of happiness" (Mill, 1861/2003, p.194). Apparently making things worse, Mill tells us that the sacrifice itself Utilitarianism doesn't consider 'good' (Mill, 1861/2003, p.194). Let me help Mill a little: the act which constitutes 'moral sacrifice' (which it does only because it counts as a *morally good act* which incidentally involves 'self-sacrifice' as a by-product of *subordination*, i.e. of putting self-love second with respect to moral duty in terms of one's practical priorities) is of itself *morally good*, i.e. it's moral goodness (i.e. its morality or moral

worth and therefore the person's being worthy of gratitude for her act) does not depend on felicific consequences or tendencies in any form. Now, of course self-sacrifice if not of itself a benefit for anyone, and of course the person that 'sacrifices her good' does not see the sacrifice itself (the act regarded as an 'act of sacrifice') as something worth pursuing. Kant is right in thinking that she does so *from duty* (simply because she *morally ought* to): Mill is right in thinking that this same act (acting from duty) essentially involves acting for the happiness of subjects, such that there is nothing that one does from duty, and therefore no 'moral sacrifice', that is not done *for the happiness of subjects*. Hence my (admittedly reconstructed) interpretation of the claim that Mill is advancing here. Again: he has reclaimed for Utilitarianism the right to demand 'self-sacrifice' no less than the Stoic or the Transcendentalist (a.k.a. Kant) are capable of demanding, he is only adding that none of this would make any sense (i.e. we would not really identify it as part of *morality*) if it didn't relate – at least at some level, and of course always in the right way – to sentient beings in general as intended beneficiaries. In my (admittedly reconstructed) reading of Kant's moral philosophy as a whole, he agrees. In virtue of Kant's 'doctrine of morally obligatory ends' (MM 6:379–395) and other elements of his morally philosophy, he already seems to agree on the fundamental point. Once again, therefore, I conclude that there is no real conflict between Utilitarianism and Kantian ethics. Neither has a claim to being more (duly) 'Stoic' than the other.

As to Kant's arbitration between Stoic and Epicurean philosophy (with which I end), there are two important points to note. (1) It is a notorious fact that Kant rejects Stoic philosophy on the basis of what he sees as the truth in Epicurean philosophy. We have gone over this more than once, but I really need to recapitulate. The Stoics are right to emphasise moral virtue as the paramount 'good' (the 'supreme practical good' as Kant would have it) in human life, but they are wrong in thinking that happiness (well-being) is thereby secured, or that everything is therefore fine in a human life; they are right in claiming that human beings are capable of something 'higher' than pleasure, but they are wrong in denying pleasure as such the claim that it has as an essential component of 'the human good': there is an irreducibly felicific component to the 'human good', and it depends completely on the presence and absence of pleasure and pain, enjoyment and suffering (in the literal, experiential sense of these terms). "Well-being or ill-being", Kant

tells us, “always signifies only a reference to our state of agreeableness or disagreeableness, of gratification or pain, and if we desire or avoid an object on this account we do so only insofar as it is referred to our sensibility and to the feeling of pleasure or displeasure it causes.” (CPrR 5:60). If it is happiness or well-being that we are talking about, Kant is as much an *unqualified hedonist* as Bentham was, as we can appreciate in his larger treatment of the issue in the *Second Critique* (his account of non-moral psychology in particular, which we find condensed in 5:23–25). And so the conclusion follows once again – now from another perspective – that there is no real conflict between Kantian ethics and Utilitarianism: neither has a claim to being more (duly) ‘Epicurean’ than the other.

Now, the second (more interesting point) is what Kant gives to the Stoics and how he tries to put it together with the ‘Epicurean’ component of his view. We have visited this place already, but (again) I need it to recapitulate and offer my final reflections. The most essential way in which Kant puts Stoic and Epicurean philosophy together is one that transcends the limits of moral or ethical theory: the most general way in which Kant does this (through his notion of the ‘highest good’ in the person) is one that pertains first and foremost to his theory of ‘the human good’, and second (and more importantly for him in terms of the general emphasis that he gives to everything he touches) in his ‘practical philosophy’, which is based on the dualistic psychology of the human being: the idea that the human being is essentially driven by both duty and self-love. Now, there is a basic thesis in relation to Kant’s dualism that does pertain in ethical theory, namely the part that says that *moral goodness* (the morally good maxim) consists in the *subordination* of the influence – in the human will – of self-love to the influence of moral duty. However, as you can see, this is Kant’s distinctive (ubiquitous to his practical philosophy) take on the *priorities* that reason attaches to the two basic principles of the human will (through his conception of the *kind of principle* that the moral principle is, of course: a *categorical* principle), which together comprise ‘practical’ as opposed to just ‘moral’ philosophy. ‘Practical philosophy’ includes both ‘moral’ and ‘non-moral’ practical philosophy, a theory of ‘pure practical reason’ that Kant associates with the Stoics and a theory of ‘empirical practical reason’ that he associates with Epicureans: he sees each school as giving us (with their own emphasis) an irreducible aspect of our ‘human – including agential – nature’, and we are supposed to

put them together in the idea of the ‘highest good in the person’, which Kant then associates with ‘Christian philosophy’.

Now, it is here that a certain issue arises on which I haven’t dared to pronounce, namely Kant’s idea of ‘retributive’ or ‘punitive’ (and also of course ‘remunerative’) justice: the idea that the virtuous should be happy and the vicious suffer. Perhaps to your disappointment, I will continue in silence with respect to my take on this particular notion, and instead make a basic point that I think needs to be made concerning discussion of it in moral philosophy. The issue of ‘punitive’ / ‘remunerative’ justice comes up very frequently in discussions surrounding the relation between Kantian ethics and Utilitarianism, much more so than *self-givingness* does (which in fact hardly ever does because Kantians are generally happy to avoid it and Utilitarians tend to treat facts about the nature of moral motivation as less than morally fundamental, most of the time not even caring what it is that moves human beings in action so long as they ‘do the right – felicifically optimific – thing’), and I want to take the chance to explain why this tendency should be regarded as seriously problematic. This tendency has to change because self-givingness is part of morality in a way that ‘punitive’ / ‘remunerative’ justice can never hope to be (even if we agree on calling morality itself by the name – as one among its various names – of ‘worthiness of happiness’). What I mean here is that the latter idea (whatever we make of it) *presupposes* the former as itself part of morality in a more essential manner, being exactly the kind of thing one can point to as rendering human beings ‘worthy of happiness’. Accordingly, we can imagine God’s judgement proceeding as follows: did she exercise adequate self-givingness? She is worthy of happiness, then.

What I just said is by no means taking sides for Utilitarianism as against Kantian ethics, for I think that a careful reading of Kant himself reveals that we are in agreement here. Kant’s own doctrine of morally obligatory ends (which puts the human being in relation to the happiness of subjects as intended beneficiaries), for instance, belongs in his moral philosophy in a way that his doctrine of ‘retributive’ / ‘remunerative’ justice can never hope to. The latter doctrine transcends ethical theory as such even for Kant, having more to do with the related question of *hope* taking us to what Kant and other people would call ‘religion’ (something which I haven’t touched directly in this work). The pursuit of everyone’s happiness in practical love, on the contrary, which essentially involves *acting for*

the benefit of subjects, is essentially constitutive of moral activity *itself*, and as such it should figure in moral philosophy well before we get to the point (if we even need to go there) of talking about punishment and reward. In this way, everyone's happiness – including the happiness of non-human animals – is more central (radically more pressing talking about) than one's morally deserved happiness in Kantian *ethical* theory.

Kant's account of the relations between moral virtue and happiness is far more complex and exhaustive – and to a large extent more 'Utilitarian', and in any case more useful for Utilitarians in their own theorising and for everyone else defending the moral standing of non-human animals – than popular slogans and narratives have ever managed to convey. Kant's most basic point in his moral philosophy (a point emphasis on which I think deserves him the label of 'the philosopher of practical priorities') is of course that morality (holding our moral principles sacred and religiously observing them) is radically more important than – in the sense of *practically unconditional* with respect to – the pursuit of our own happiness in self-love: morality is always to be given 'first place', never to be compromised. But of course this point (which should on its own be appealing for Utilitarians as philosophers who claim to give morality a prominent place in human life) cuts across other themes in Kant's philosophy where happiness figures in a more prominent manner: there is also Kant's 'doctrine of morally obligatory ends' (MM 6:379–395), his derivation of the duty of generalised benevolence entailing the notion of 'self-givingness' or 'acting for the happiness of subjects' (MM 6:393), and more.

I hope my use of these resources, and this work as a whole, have shed some light on the nature of or our existing ethical theories. More importantly, I hope light has been shed on what it means to be a human being.

Bibliography

- Altman, M.C. 2019. Animal Suffering and Moral Salience: A Defense of Kant's Indirect View. *Journal of Value Inquiry*. 53 (2), pp.275-288.
- Bader, 2015. Kantian Axiology and the Dualism of Practical Reason. In: Hirose, I. & Olson, J. eds. *The Oxford Handbook of Value Theory*. Oxford, UK: Oxford University Press, pp. 175-201.
- Bentham, J. 1907. *An Introduction to the Principles of Morals and Legislation*. Oxford, UK: Clarendon Press.
- Cavalieri, P. 2001. *The Animal Question: Why Nonhuman Animals Deserve Human Rights*. New York, NY: Oxford University Press.
- Crisp, R. 1997. *Mill on Utilitarianism*. London, UK: Routledge.
- Cummiskey, D. 1996. *Kantian Consequentialism*. New York, NY: Oxford University Press.
- Darwall, S. 1977. Two Kinds of Respect. *Ethics*. 88 (1), pp.36-49.
- Darwall, S. 2006. *The Second-Person Standpoint: Morality, Respect and Accountability*. Cambridge, MA: Harvard University Press.
- Darwall, S. 2012. Bipolar Obligation. In: Shafer-Landau, R. ed. *Oxford Studies in Metaethics Volume 7*. Oxford, UK: Oxford University Press, pp.333-357.
- Dean, R. 2006. *The Value of Humanity In Kant's Moral Theory*. Oxford, UK: Oxford University Press.
- Denis, L. 2000. Kant's Conception of Duties regarding Animals: Reconstruction and Reconsideration. *History of Philosophical Quarterly*. 17 (4), pp.405-423.
- Driver, J. 2012. *Consequentialism*. Oxford, UK: Routledge.

- Fletcher, J. 1966. *Situation Ethics: The New Morality*. Louisville, KY: Westminster John Knox Press
- Frankena, W. 1973. *Ethics: Second Edition*. Englewood Cliffs, NJ: Prentice-Hall Inc.
- Frankena, W. 1976. Love and Principle in Christian Ethics. In: Goodpaster, K.E. ed. *Perspectives on Morality: Essays by William Frankena*. Notre Dame, IN: University of Notre Dame Press.
- Gruen, L. 2011. *Ethics and Animals: An Introduction*. Cambridge, UK: Cambridge University Press.
- Kagan, S. 1992. The Structure of Normative Ethics. *Philosophical Perspectives*. 6, pp.223-242.
- Kant, I. 1996. *Practical Philosophy, The Cambridge Edition of the Works of Immanuel Kant*, Gregor, M. ed. and trans. Cambridge, UK: Cambridge University Press.
- Kant, I. 1998a. *Religion within the Boundaries of Mere Reason*. Cambridge, UK: Cambridge University Press.
- Kant, I. 1998b. *Critique of Pure Reason, The Cambridge Edition of the Works of Immanuel Kant*. Guyer, P. and Wood, A. W. eds. and trans. Cambridge, UK: Cambridge University Press.
- Kant, I. 2001. *Lectures on Ethics*. Heath, P. ed. and trans. and Schneewind, J. B. ed. Cambridge, UK: Cambridge University Press.
- Kant, I. 2006. *Anthropology from a Pragmatic Point of View*. Lauden, R. B. (ed. and trans.). Cambridge, UK: Cambridge University Press.
- Kant, I. 2007. Conjectural Beginning of Human History. Wood, A. W. (trans.). *Anthropology, History and Education*. Cambridge, UK: Cambridge University Press, pp. 160-175.

- Kant, I. 2011. *Groundwork of the Metaphysics of Morals: A German-English Edition*. Gregor, M. and Timmermann, J. eds. and trans. Cambridge, UK: Cambridge University Press.
- Kain, P. 2010. Duties regarding Animals. In: Denis, L. ed. *Kant's Metaphysics of Morals: A Critical Guide*. Cambridge, UK: Cambridge University Press.
- Lautrup, J. 2015. *From Duty, For Happiness – Authority and Value in Kant's Theory of Practical Reason*. Ph.D. thesis, University of Oslo.
- Mill, J. S. 2003. Utilitarianism. In: Warnock, M. ed. *Utilitarianism and On Liberty*. Second Edition. Malden, MA: Blackwell Publishing Ltd., pp. 181-235.
- Mulgan, T. 2006. *Future People: A Moderate Consequentialist Account of our Obligations to Future Generations*. Oxford: Oxford University Press.
- Mulgan, T. 2007. *Understanding Utilitarianism*. Stocksfield, UK: Acumen Publishing Limited.
- Rawls, J. 1999. *A Theory of Justice: Revised Edition*. Cambridge, MA: Harvard University Press.
- Rousseau, J. J. 1987. Discourse on the Origin of Inequality. In: Cress D.A. ed. *Basic Political Writings of Jean-Jacques Rousseau*. Indianapolis, IN: Hackett Publishing Company.
- Rowlands, M. 2009. *Animal Rights* (2nd Ed.). Miami, FL: Palgrave Macmillan.
- Scanlon, T. M. 2000. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Scarre, G. 1996. *Utilitarianism*. London, UK: Routledge.
- Scheffler, S. 1982. *The Rejection of Consequentialism*. New York, NY: Oxford University Press.

Sensen, O. 2011. *Kant on human dignity*. Göttingen: De Gruyter.

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

Singer, P. 2011a. *Practical Ethics* (3rd Ed.). New York: Cambridge University Press.

Singer, P. 2011b. *The Expanding Circle*. Princeton, NJ: Princeton University Press.

Singer, P. 2017. *Utilitarianism: A Very Short Introduction*. Oxford, UK: Oxford University Press.

Sinnott-Armstrong, W. 2019. Consequentialism. [Online]. The Stanford Encyclopedia of Philosophy. Available from:

<https://plato.stanford.edu/archives/sum2019/entries/consequentialism/>

Smith, A. 1976. *The Theory of Moral Sentiments*. Oxford, UK: Oxford University Press.

Timmermann, J. 2005. When the Tail Wags the Dog: Animal Welfare and Indirect Duty in Kantian Ethics. *Kantian Review*. 10, pp.128-149.

Wood, A. W. and O'Neill, O. 1998. Kant on Duties Regarding Nonrational Nature. *Aristotelian Society Supplementary Volumes*. 72 (1), pp.211-228.