1. Introduction

Human rights are rights that humans have simply by virtue of being human. Human rights are universal, they apply to all humans, and categorical, their validity does not depend on the contingent preferences, beliefs or desires of (groups of) individuals. How could we possibly justify such universal and categorical rights? How can the idea of human rights be vindicated? This question is all the more pressing given the existence of widespread skepticism about the universal validity of human rights, both in society and in academia (see e.g. the contribution to this panel by Dascha Dürring).

The assumption of this paper is that human rights are not (just) legal rights, but (also) moral rights. Consequently, I will discuss the question of the appropriate method of moral justification of human rights. Given the limitations of this short paper, I will not discuss the justification of specific human rights nor their practical implementation (this task is taken up, in different ways, by both Jurriën Hamer and Gerhard Bos).

Within the scope of this short paper, I also cannot go into the details of different methods of the justification of human rights. Instead, the aim of this paper is merely to provide a brief overview of different methods and to provide at least some suggestions as to why we should prefer a transcendental justification of human rights. I will proceed as follows: first, I will sketch the outline of a so-called ‘transcendental’ justification of human rights by building on the work of Alan Gewirth. Second, I will briefly argue why a transcendental justification should be preferred to alternative methods of moral justification – I will focus specifically on so-called teleological justifications and the method of reflective equilibrium.

2. A Transcendental Justification of Human Rights

A transcendental argument is an argument which tries to show that there are certain conclusions that one cannot coherently deny from one’s own first-person perspective
as an agent, i.e. as someone who wants to achieve certain ends in the world (e.g. to
drink coffee, to pursue a career, to get married, etc). A transcendental argument for
human rights thus tries to show that a commitment to human rights follows from
being an agent. Although there are different transcendental approaches to the
justification of human rights, I focus exclusively on a Gewirthian transcendental
argument which tries to show that any agent necessarily has to accept that all agents
(including herself) have a right to whatever are the necessary preconditions of agency
(Gewirth 1978; Gewirth 1992; Beyleveld 1991). These necessary preconditions of
agency include a certain measure of freedom and a certain amount of physical and
psychological well-being.

One way to present the Gewirthian, transcendental argument for human rights
is as follows: First, the argument starts from the observation that we understand
ourselves as agents, i.e. as beings which try to achieve certain purposes. For instance,
we aim to buy coffee, read books, have drinks, etc. So much seems to be
uncontroversial. Second, it is argued that if I (or any other agent) want to pursue a
purpose or an end, I ought to take the necessary means to that end or give up the end
(this is referred to as the instrumental principle). For instance, if I want something to
drink, and there is only water available, then I ought to drink water. Of course, once I
discover that there is only water available, I might revise my end and decide that I
would rather be thirsty than drink water, but if I choose to stick to the initial goal, then
I ought to drink the water. Someone who wants an end but who denies that she
ought to take the means to an end (or give up the end) simply fails to understand what is
involved in wanting an end and this person therefore becomes un-intelligible to both
herself and to others.

The example of having to take the necessary means to my end of getting a
drink is contingent on wanting to a specific end. Only someone who wants a drink,
ought to drink the water. Clearly this is a far cry from justifying human rights. The
transcendental approach to the justification of human rights, however, uses this basic
model of (instrumental) normativity to argue for human rights. The remainder of the
argument proceeds as follows:

---

1 Other transcendental arguments start from our interaction in certain practices (Apel 1980;
O'Neill 1990; Habermas 1990; Darwall 2006). I think these arguments fail to justify categorical
and universal rights but I cannot discuss that in this paper.
In a third step, it is argued that there are not only instrumental oughts that are conditional on the acceptance of a contingent end, but that there are also instrumental oughts involved in adopting *any* end whatsoever. For instance, being alive is a necessary means for engaging in any action whatsoever. Independent of whether one wants a drink, wants to play chess, do philosophy or start a family, a necessary means for acting, whatever one’s contingent ends, is that one is alive. Even someone who commits suicide has to be alive in order to achieve his or her end (cf. Gewirth 1978, 136–137). In this sense, one ought to want to be alive, independent of one’s contingent ends. Gewirth calls these necessary means to engage in any action whatsoever the ‘generic means of agency’ (which can be contrasted with the contingent means of a specific action).

Fourth, Gewirth argues that the claim that one ought to want to have the generic means of agency (e.g. to be alive) implies that you ought to want others to not interfere with your generic means of agency, which is equivalent to saying that you ought to claim a right to the generic features of agency. However, that you ought to claim a right to the generic features of agency does not yet show that others have to respect this right or that you have to respect the rights of other agents. In order to argue for the latter conclusion, Gewirth puts forward ‘the argument from the sufficiency of agency,’ which is arguably the most controversial step of the transcendental approach. This argument states, roughly, that given that agency is a sufficient condition for claiming a right to the necessary means of agency, one ought not only to claim a right to the necessary means of one’s own agency, but one ought also to respect the necessary means of the agency of other agents - on pain of denying that agency is a sufficient condition for claiming a right to the necessary means of agency (Gewirth 1978, 101). If this argument is successful it thus follows that any agent ought to accept that all other agents have a right to the generic means of agency.

Much more need to be said about this argument, both in terms of its justification but also about its substantive implications (for the latter see the paper in this panel by Jurriën Hamer). Let me, however, finish the exposition of the Gewirthian argument for human rights by highlighting its main features. First, and most importantly, the argument proceeds from the first person. The argument tries to show what anyone who understands herself as trying to achieve purposes is necessarily committed to. This first-person perspective is sufficient for the justification of human rights, because if the argument is successful it shows that any
agent necessarily has to accept the rights to the generic means of agency of all other agents (including herself).

Second, on this approach the source of the normativity of human rights is no more mysterious than the source of normativity of instrumental reasons. The only difference between contingent instrumental means and the generic means of agency is that whereas contingent means should only be accepted insofar as someone wants a specific end, generic means should be accepted insofar as there are beings who want to achieve any end whatsoever. Of course, one might deny that there are any generic means of agency, or object to the argument from the sufficiency of agency, but one of the benefits of the argument is that it can explain how human rights are continuous with other domains of practical reasoning.

3. Other Approaches to the Justification of Human Rights

Instead of further elaborating the transcendental approach to human rights, I want to briefly indicate why the transcendental approach should be preferred to alternative approaches to moral justification. I will focus on two dominant alternatives 1) a teleological justification of human rights. 2) the justification of human rights through the method of reflective equilibrium.

3.1 Teleological justifications

Consider, first, a teleological justification of human rights. A teleological justification of human rights aims to ground human rights in certain pre-existing values. James Griffin, for instance, argues that personhood has (intrinsic) value and that human rights should consequently understood in terms of the protection and development of personhood (Griffin 2008, 57–58).

The problem with the teleological approach is that it simply starts from an allegedly self-evident value, which is both ontologically (what is the ontological status of this value; where and how does it exist?) and epistemologically (when and for whom is it self-evident?) problematic. This does not mean that it is impossible to account for the value of personhood, but that there is at least a prima facie reason to prefer a moral methodology without the epistemological and ontological burdens of teleological justifications of human rights.
The benefit of a transcendental argument is that it does not presuppose the value of personhood of agency. Instead, the argument tries to show that anyone who understands him or herself as an agent is rationally committed to accept human rights to the generic means of agency. A transcendental argument therefore neither presupposes some kind of mind-independent value, nor claims that certain values are simply self-evident.

3.2 Reflective equilibrium

The most dominant approach to moral justification in general is the method of reflective equilibrium, a coherentist method of justification according to which a moral judgment is justified if it coheres with all considerations that are relevant to the issue at hand. These relevant considerations typically include initial judgments about cases, moral principles, and relevant background theories (Daniels 1979). This method is also employed to justify human rights. In The Law of Peoples, Rawls, for instance, claims that (universal and categorical) human rights can be justified in reflective equilibrium (2000, 86f32; For similar approaches see Liao 2015; Gilabert 2015; Nussbaum 2001, 101–103).

The main advantage of the method of reflective equilibrium is that, unlike teleological justifications, it does not presuppose any (self-evident) values. The problem with the method of reflective equilibrium, however, is that it is unclear how this methodology could lead to the justification of any universal and categorical moral judgments, including judgments about human rights. For instance, it is unclear how the method of reflective equilibrium can convince a skeptic about human rights that she should accept the universality of human rights, given that a skeptic simply does not share the initial judgments of the proponents of human rights. One could modify the method of reflective equilibrium in such as way as to simply exclude certain positions (for instance by proposing a conception of ‘considered judgments’), but this would undermine reflective equilibrium as a distinctive coherentist methodology. It seems at best implausible that a method that starts from contingent (moral) judgments of individuals can lead to categorical and universal conclusions (see De Maagt, forthcoming).

4. Conclusion
A transcendental approach seems to combine the best of both worlds: it can agree with defenders of reflective equilibrium that there are no self-evident moral values on which to base human rights, while it diverges from reflective equilibrium by focusing not on the coherence of our contingent belief set, but on those judgments any agent necessarily has to accept. A transcendental argument agrees with teleological approaches to justification that there are universal and categorical universal rights, but it denies that these can only be grounded in pre-existing values that are known self-evidently.

If there is a rational basis for universal and categorical human rights, a transcendental argument thus seems to be our best hope of providing one.