

Lund university
Department of Philosophy
Supervisor: Dan Egonsson

Christoffer Ivarsson
870425
Bachelor thesis

Reasonable Contractors

2013-01-14

References and Abbreviations

To make things a bit easier for the reader I have gathered a list of abbreviations that are frequently used in the paper.

Concepts

OP: Original Position

VI: Veil of Ignorance

FEO: Fair Equality of Opportunity

DP: Difference Principle

Books and articles

TJ / *Theory*: *A Theory of Justice (revised edition)*, Rawls, John; Harvard University Press, 1999

RR: *Reconstructing Rawls – the Kantian foundations of justice as fairness*, Taylor, Robert S. Pennsylvania State University Press, 2011

PL: Political liberalism, Rawls, John; Columbia University Press, New York. 1993

GMM: *Groundwork of the Metaphysics of Morals*, Kant, Immanuel; edited by Mary Gregor and Jens Timmermann, Cambridge University Press, 2012

Except for the encyclopedia entries and Alexander's and Arneson's articles, references will be done by writing out the entirety of the title the first time, and subsequently the abbreviations will be used, so for instance “(RR p. 56)” means that page 56 in *Reconstructing Rawls – the Kantian foundations of justice as fairness* by Robert S. Taylor's is being referred to.

Contents

| | |
|--|----|
| References and Abbreviations..... | 2 |
| Contents..... | 3 |
| Introduction..... | 4 |
| Part 1 | |
| 1 Justice as Fairness..... | 5 |
| 1.1 A brief overview of how the Fair Equality of Opportunity principle fits into Justice as Fairness..... | 7 |
| 1.2 The Critique..... | 9 |
| 1.2.1 The Charges from Alexander..... | 10 |
| 1.2.2 Arneson's Critique..... | 12 |
| Part 2 | |
| 2 Robert S. Taylor's reconstructed Theory of Justice..... | 14 |
| 2.1 Highest and higher-order interests becomes Moral and personal autonomy..... | 15 |
| a) Reasonableness – Moral Autonomy..... | 15 |
| b) Rationality – Personal Autonomy..... | 17 |
| c) Derivative interests..... | 18 |
| d) Where is he going?..... | 19 |
| e) Adding Self-realisation..... | 20 |
| 2.2 The Principles of Justice and their priority..... | 21 |
| a) The interests in moral and personal autonomy yields the liberty principle..... | 22 |
| b) The interest in self-realization yields the Fair Equality of Opportunity principle..... | 23 |
| Part 3 | |
| 3 Problems and advantages of Taylor's rendition of the theory..... | 27 |
| a) The black hole-argument – is it a red herring?..... | 27 |
| b) The second answer: The threshold..... | 28 |
| 3.1 Some specific objections..... | 32 |
| Conclusion..... | 34 |

Introduction

In this paper I will examine Robert S. Taylor's reconstruction of John Rawls' *Justice as Fairness*. The paper will mainly be focused around the Fair Equality of Opportunity principle and how it is defended in Taylor's rendition of *Justice as Fairness*. The reason I chose this approach to Taylor's theory is that he himself starts out in his 2005 article "*Self-Realization and the Priority of Fair Equality of Opportunity*" by saying that the Fair Equality of Opportunity principle could be saved from much of the critique that has been raised against it if it were understood in light of the interest in self-realisation. However, we don't get a satisfactory explanation of how the interest in self-realisation is supposed to be understood until his 2011 book entitled "*Reconstructing Rawls – The Kantian foundations of Justice as Fairness*" which sets the interest in self-realisation within a larger framework of a completely reconstructed *Justice as Fairness*.

I begin by giving a brief overview of *Justice as Fairness* as it is generally understood, and then describe how the Fair Equality of Opportunity principle fits into *Justice as Fairness* before looking at the critique raised against it that Taylor claims he can defend against. From there I proceed to examine Taylor's reconstructed theory. I look specifically at the basis for his reconstruction, and I evaluate and discuss the merits of his defences of the FEO principle. I will argue that the way this defence is presented is far from optimal and bears signs of an *ad hoc*-defence.

It will be noted however, that while there is problems with Taylor's theory, those problems that I point out are more a result of Taylor's attempts to stay close to Rawl's original theory, and is perhaps not too damning for the basis of his theory. It should be noted and underlined that what I present here is not a comparison of Taylor and Rawls, and I do not purport to be able to make a proper evaluation as to the feasibility of Taylor's theory being the road that Rawls could have taken. My concern here is Taylor's theory and what promises it may or may not hold.

Part 1

1. Justice as Fairness

Perhaps the most striking and distinctive part of John Rawls' theory (called *Justice as Fairness*, presented in the book entitled *A Theory of Justice*) is the thought-experiment described as the original position (henceforth abbreviated OP). This is the imagined contract situation in which we are to imagine choosing the principles of justice that make up the substantive part of *Justice as Fairness*; those principles that are to regulate and determine the basic social structure of society. What makes Rawls' imagined contract situation different from earlier imagined contract situations, for example Hobbesian and Lockean ones, is how we are to understand the constitution of the contractors in the original position. (Freeman, Samuel, "*Original Position*", The Stanford Encyclopedia of Philosophy (Spring 2012 Edition), Edward N. Zalta (ed.))

Most notably, Rawls places the contractors in the original position behind a *Veil of Ignorance* (henceforth abbreviated VI), which is a feature designed to guarantee that the principles the contractors choose are indeed fair (hence the name *Justice as Fairness*). The veil of ignorance therefore robs the contractors of specific information about who they are. When in the OP, the contractors don't know their age, their social status, how smart, strong or skilled they might or might not be; in short, they know nothing about themselves or what they might or might not have in terms of assets; natural or social. Added to this, the veil of ignorance also robs them of knowledge about their "conceptions of the good" and "their special psychological propensities" - which is to say, they don't know specifically what they want in life. (Freeman, Samuel, "*Original Position*", The Stanford Encyclopedia of Philosophy (Spring 2012 Edition), Edward N. Zalta (ed.))

Of course, the original position cannot merely reside in a few utterly oblivious individuals not knowing anything, they are there to choose principles; to decide the just ordering of society, and so they must have some basis for choosing between different possible principles. And the contractors do have access to a lot of general information behind the veil, such as for example a general knowledge about economics, psychology, political science, and biology and other natural sciences. They are said to have knowledge about "the general tendencies of human behaviour and psychological development, about biological evolution, and about how economic markets work, including neo-classical price theory of supply and demand." and most importantly they know about the so called circumstances of justice – the fact that people aren't always overly altruistic combined with the fact that the world isn't a place of abundance. So the contractors know that people generally desire "primary social goods" - which is a generic term for all those things that people

generally need to be able to live a good life and develop their “moral powers”. The principles of justice are chosen to ensure the fair distribution of these goods. (Freeman, Samuel, *“Original Position”*, The Stanford Encyclopedia of Philosophy (Spring 2012 Edition), Edward N. Zalta (ed.))

The “moral powers” mentioned above is part of what *motivates* the contractors in the original position. Following the familiar idea that we need something more than knowledge alone to motivate us to actually choose something, it is supposed that the contractors have some motivations – or, as it will later be referred to by Taylor, “regulative interests” - that drive them to choose the principles of justice. (Taylor, Robert S, *Reconstructing Rawls – the Kantian foundations of justice as fairness*, Pennsylvania State University Press, 2011, p. 22-3)

According to Rawls the parties in the original position have two moral powers, those of being *rational* and *reasonable* (*having a sense of justice*). Being rational means having a *conception of the good* and being able to form a *rational plan of life* aimed at a subjective conception of the good. Simply put, the contractors are rational, in the sense that they know that they are both able to, and motivated towards, formulating a rational plan of life aimed at fulfilling their conception of the good - whatever that may be – and therefore they choose principles of justice that guarantee them the means to do this. These means are what is called social primary goods and are described as “rights and liberties, powers and opportunities, income and wealth, and the social bases of self-respect”. The fact that the contractors are conceived of as having a sense of justice (a capacity for reasonableness) limits and frames the rational plan of life. Whereas rationality concerns the contractors’ conception of the good, their sense of justice, or reasonableness, concerns their concept of right, and in *Justice as Fairness* the right is prior to the good. On one reading of Rawls, the moral power of reasonableness should be understood as something that is needed for social cooperation to work; had we not a sense of justice we would be considered unreasonable, and unable to partake in social life (and subsequently be unable to effectively pursue our conception of the good). There is however another possible way of understanding the interest in developing the moral powers of rationality and reasonableness, and that will be explored when Robert S. Taylor's account of *Justice as Fairness* is addressed later on. (Freeman, Samuel, *“Original Position”*, The Stanford Encyclopedia of Philosophy (Spring 2012 Edition), Edward N. Zalta (ed.))

Rawls uses the idea of the original position together with the idea of reflective equilibrium. We should work from both ends in searching for the principles of justice, he writes, considering both what the contractors in the original position would choose and balancing their convictions against the firmly held convictions of justice we already have. Some convictions will be revised in light of what the original position situation yields, but sometimes we have firmly held convictions that we would not want to revise, and so this must be accounted for in the original position. Rawls

notes as an example that we are confident that religious intolerance and racial discrimination are unjust, and that these are convictions of justice that we would not want to revise. In going back and forth between convictions, and revising them where it is needed we can eventually reach reflective equilibrium, where we have an initial situation (the OP) that yields principles that are reasonable and sufficiently match our considered judgements. The care that Rawls takes to not only base the principles of justice on what rational and reasonable contractors in the original position would choose, but also to balance that against our actual considered judgements is important to note, because it stands in contrast to Taylor's reconstruction. Rawls writes, after describing the process of reflective equilibrium, that the principles of justice cannot be derived from necessary truths, and in order to justify a specific conception of justice we must accommodate both "reasonable philosophical conditions" (i.e. what the OP agents would choose) and "considered judgements of justice" (roughly; what we actually think). Taylor's account, as we will see, stands in stark contrast to Rawls' conviction that principles of justice cannot be derived from necessary truths. (Rawls, John; *A Theory of Justice (revised edition)*, Harvard University Press, 1999, p 17-19)

I turn now to the principles of justice that are chosen by the contractors in the original position (checked against the process of revision of considered convictions resulting in reflective equilibrium) and in particular a sub-component of one of those principles that is the starting point of this paper: the *fair equality of opportunity principle*. In the course of trying to defend this principle Robert S. Taylor ends up reworking the basis for *Justice as Fairness*, namely the contractors interests in rationality and reasonableness. Before I get to that though, I will briefly present the critique raised against the fair equality of opportunity principle, which will give me the opportunity to evaluate the merit of the proposed defence.

1.1 A brief overview of how the Fair Equality of Opportunity principle fits into *Justice as Fairness*

The rational contractors choose principles behind the veil of ignorance that will govern society, or provide a "just arrangement of the major political and social institutions of a liberal society: the political constitution, the legal system, the economy, the family, and so on."(Wenar, Leif, "*John Rawls*", The Stanford Encyclopedia of Philosophy (Winter 2012 Edition), Edward N. Zalta(ed.))

The principles that emerge from this original contract situation are twofold. There is the **first principle** which we may call the *fair liberty principle* and there is the **second principle** which is of course two principles, namely the *Fair Equality of Opportunity principle (henceforth abbreviated FEO)* and the *Difference Principle (henceforth abbreviated DP)*:

“First Principle: Each person has the same infeasible claim to a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all;

Second Principle: Social and economic inequalities are to satisfy two conditions:

- a. They are to be attached to offices and positions open to all under conditions of fair equality of opportunity;
- b. They are to be to the greatest benefit of the least-advantaged members of society (the difference principle). (JF, 42–43)” (Wenar, Leif, "John Rawls", The Stanford Encyclopedia of Philosophy)

These principles are strictly ordered, in fact, they are as strictly ordered as principles can be. The First principle, the fair liberty principle enjoys what Rawls calls lexical priority over the second principle, and the FEO principle likewise enjoys lexical priority over the DP principle.

What that essentially means is that if a principle takes lexical priority over another, there may never be trade-offs from higher to lower. In particular, the political liberties covered by the first principle may never be restricted in order for society to attain (for example) greater wealth for its citizens. And since the FEO principle is lexically prior to the difference principle, we may never reserve certain positions and offices for certain groups of people in order to increase the benefits that can be distributed via the difference principle. This last part in particular is very sketchily described, and as we will see, not entirely correct, but since it is the description given when Rawls is being introduced to the new philosophy student, it gives us a good starting point from which to work. (Wolff, Jonathan *“An introduction to political philosophy”* Oxford University Press, New York 2006, p. 158)

The fair equality of opportunity principle is described by Rawls as follows:

“The thought here is that positions are to be not only open in a formal sense, but that all should have a fair chance to attain them. Offhand it is not clear what is meant, but we might say that those with similar abilities and skills should have similar life chances. More specifically, assuming that there is a distribution of natural assets, those who are at the same level of talent and ability, and have the same willingness to use them, should have the same prospects of success regardless of their initial place in the social system. In all sectors of society there should be roughly equal prospects of culture and achievement for everyone similarly motivated and endowed. The expectations of those with the same abilities and

aspirations should not be affected by their social class” (TJ p. 63)

As the quote clearly shows, the FEO principle can be described as a kind of anti-discrimination principle which is supposed to correct for the arbitrariness of which social class one is born into. It can however be interpreted as entailing both more and less; as being both too strong and too weak as a anti-discrimination principle. I will now turn to the critique raised against the principle.

1.2 The Critique

Even though the critique and discussions concerning the fair equality of opportunity principle is no doubt extensive, for the purpose of this paper I will limit myself to the critique addressed in Taylor's above mentioned article. The critique in question is raised by Larry Alexander, in his rather famous article from 1986 “Fair Equality of Opportunity: John Rawls (Best) Forgotten Principle” as well as by Richard J. Arneson in his 1999 article “Against Rawlsian Equality of Opportunity”.

I am thankful to Eric Brandstedt for pointing out to me that much of the critique presented here could be charged with missing the mark. This would be because both Arneson and Alexander are talking about positions that do not belong to the basic social structure that *Justice as Fairness* and the FEO principle concerns itself with. However, when Rawls describes what fair equality of opportunity substantially means, he says that the government is to ensure equal chances of occupation (which can be done either by having a public school system or subsidizing private schools) but it is also supposed to enforce equality of opportunity in economic activities and the free choice of occupation – which is done by “policing the conduct of firms and private associations” as well as making sure that desirable positions aren't monopolised somehow. (TJ, p 243)

This could be seen as constituting a deviation from the main subject of *Justice as Fairness* as being that of the basic social structure. I myself am of the opinion that talk of fair opportunity would become rather meaningless if the things talked about above weren't included, and I think that major social and economic institutions (which apart from political and legal institutions constitute the basic structure) could be broadened to include pretty much anything. The family, for instance, is said to be an example of a major social institution, as I think it must be. In any case, Taylor interprets the FEO principle as having as it's subject positions in “governments, private and public corporations, universities, NGOs, and so forth”, which I would argue is quite reasonable on his reinterpretation of the basis of the principles of justice. (RR p. 180)

This means that the criticism raised against the FEO principle must be handled in some way by Taylor. So, I will proceed to present the criticism, although since a lot of the criticism that is

raised in the articles doesn't really hit Taylor's reconstructed version of *Justice as Fairness* it will be a brief review.

1.2.1 The Charges from Alexander

Larry Alexander argues that the lexical priority of FEO is not only intuitively unattractive, but also that it shows an inconsistency in Rawls' method of deriving principles from what the rational contractors in the original position (OP) would choose.

I will present some of the different examples he uses to show the implausibility of the contractors choosing the principle and making it lexically prior to the difference principle. Generally the critique can be characterised as a somewhat standard critique from a consequentialist standpoint. (Alexander, Larry "Fair Equality of Opportunity: John Rawls (Best) Forgotten Principle" *Philosophy Research Archives* 11 1986 p. 197-208)

If we are to understand the priority of FEO as the absolute priority to equalize the competitive positions of those with the same inborn abilities and the same willingness to use them, would we expect the contractors in the OP to choose the FEO principle with this strong priority? Alexander's constructs a few scenarios in order to show that understanding the FEO principle in this way would not lead the OP contractors to give it this priority. (Alexander, 201)

I) The Alaska Case:

Two persons, A and B, have the same abilities needed to handle a certain job. They also have an equal willingness and ambition to get the job in question. The problem is that the job is in North Dakota, and while A lives in North Dakota, B happens to reside in Alaska, and the people hiring for the job can't afford to fly B in for the interview. The question then becomes: "Because of the interviewing cost, must the government, perhaps at outrageous costs in income, see that interviewing opportunities as well as talents are equalized for those of the same inborn abilities in order to comply with Rawls' principle?"(Alexander, 201)

II) The Socialized Ambition Case

Other difficulties notwithstanding, there seems to be an inconsistency in the way FEO is supposed

to work, in that it hinges on personal ambition and willingness whether or not it is violated in a given situation. Alexander notes that Rawls goes to great lengths to devise principles that are to equalize inequalities that are “the product of undeserved endowment in the morally arbitrary natural lottery and/or undeserved social circumstances”. Now, if A and B have equal natural talents, but due to more favourable family conditions, A has developed a greater willingness to 'reach for his goals', as it were, why must society not spend resources on bringing B:s ambition level up to be on par with A:s, whatever the cost? Or, asks Alexander, why shouldn't we forbid employers to take A:s superior willingness into account? (Alexander, 201)

III) The Unequal Abilities Case

Taking the previous questions even further, given that inborn abilities are after all “as morally arbitrary for Rawls as the extent of their development” then why shouldn't society be obliged to spend whatever resources it took to equalize abilities too? And in relation to The unequal Abilities Case, Alexander notes that the extreme measures that would have to be taken to equalize natural abilities surely would seem implausible to Rawls' OP contractors, but then, he asks, why should the “half-hearted egalitarianism of the fair equality of opportunity principle” seem any more plausible? (Alexander, 201)

IV) The Wasted Resources Case

Imagine a closed system. We suppose that we have two people competing for two jobs, a skilled job and a menial job. A and B are born with the same talents, but due to circumstances, A is much better qualified for the skilled job. Rawls' FEO principle would seem to demand that society spend whatever resources necessary so that B may be on equal footing with A. So, B is educated at great expense, and A and B then have an equal 50-50 chance of getting the job they both want. Consider then, says Alexander, what happens when B gets the skilled job. A is now the worse off, and at this point, “the resources used to educate B is no longer available to compensate A for his menial job”. The point Alexander makes is that if we waste resources in this way, the worst of is still going to be worst of, but worse than could have been if the resources had been used more efficiently. (Alexander, 202)

And we arrive here at a crucial point that Alexander makes: “If, however, equality of opportunity does not translate into better jobs or income for the worst off, it is difficult to understand why Rawls' rational contractors in the original position would value it”. (Alexander,

I shall refer to this last point as the *utility argument*, and it is one I will return to as one of the main problems for the FEO principle. The cases above can be summed up with this utility argument, and with a supportive argument that I will call the *weakness argument*.

In summation, the utility argument states that since we can show that the FEO principle, if followed would demand such an unreasonable waste of resources, perhaps with little prospective gain, that the contractors in the original position surely would reject it, or at least not give it such strong priority. The weakness argument casts further doubt on the FEO principle by showing that it might not be well designed enough to be able to counteract the injustices that it was designed to prevent in the first place. Subtle forms of manipulation may still produce the same constricting effects on people.

Alexander does concede that the Aristotelian principle may provide some support for the idea that the contractors would prefer careers demanding maximum virtuosity after a certain material had been reached, but, the Aristotelian principle still wouldn't motivate the lexical priority of FEO, and in any case, Rawls doesn't provide any convincing arguments for the Aristotelian principle, nor shows that it would have this effect if correct. (Alexander, 206)

1.2.2 Arneson's Critique

Arneson's critique follows the same general, consequentialist lines as Alexander's, and so I will not mention them in the same detail. He also argues that the FEO principle seems unable to protect against "ambition formation" for instance. For him, the FEO principle should be understood as "an extension of the ideal of nondiscrimination" and as such it is both too strong and too weak.

(Arneson, Richard J, *Against Rawlsian Equality of Opportunity*, Philosophical Studies 93, No 1, 1999, p 77)

Manipulation of ambition has already been mentioned as one way in which he believes the principle is too weak as an antidiscrimination principle, and another is brought to bear when he cites Alexander (from another article) as raising the problem of taste not being ruled out as a basis for qualification – therefore, a sports team could choose to only let Caucasian players on the team, referring to the fact that the teams bigoted fans prefer that, and that if they didn't employ this practice they would loose revenue. "Wrongful discrimination" needs to be better defined, as indeed he thinks it has been done by Alexander. (Arneson, 79)

Arneson raises the "black hole-argument" that I will later return to and address at some

length, and he describes it in the following way:

“Since Fair Equality is assigned strict lexical priority, society is obligated to expend no resources at all to improve the holdings of primary social goods other than basic liberties of the worst-off members of society so long as these resources could be used instead to improve, even by the tiniest fraction, the degree to which the Fair Equality of Opportunity Principle is fulfilled” (Arneson, 81)

This leads to utterly appalling conclusions, Arneson argues. If we imagine a society where much has been done to implement Fair Equality of Opportunity, but it still remains the case that individuals with great potential born in upper middle class households still grow up with lesser prospects of success than likewise endowed individuals born in upper class households – then “Rawlsian justice demands that if by huge expenditure of resources we can offer extra education to the upper middle class youths that marginally improves their prospects of competitive success as compared to their counterparts with talents born into even more privileged households, we must do so”. (Arneson 82)

This black hole-argument is clearly one of the arguments Arneson finds most damning for the FEO principle. The resources that are drained by the FEO principle in this manner could instead be used and allocated in accordance with the Difference Principle to ensure great gains in social justice – and this trade-off simply cannot be defended on his view. He also raises the connected point that the FEO principle leaves the “severely untalented” behind, and that no resources will come their way, even though it could enhance their lives despite the fact that they have no discernible talents to develop and perfect.(Arneson, 85-6)

Part 2

2 Robert S. Taylor's reconstructed *Theory of Justice*

In his 2005 article “Self-Realization and the Priority of Fair Equality of Opportunity” Robert S. Taylor sets out to show that the priority of the FEO principle can be defended against the critique raised by Alexander and Arneson by basing it in our interest in self-realization through work. However, in my estimate the article doesn't really provide a complete argument, but only outlines a possible one. In the article we are told that the priority of the FEO principle can be defended using the Aristotelian Principle, spoken of by Rawls in *Theory*, as well as the Humboldtian Principle, likewise spoken of in *Theory*.

It is recognised in the article however, that it is necessary to ground the defence in the concept of self-realisation, which in turn needs to be grounded in the idea of autonomy. It is in his 2011 book “Reconstructing Rawls” that Taylor really lays out his theory about how Rawls theory should be re-grounded in the idea of autonomy, and it is clear that the defence of the FEO principle outlined in the 2005 article cannot stand alone, but only becomes meaningful within the full framework presented in the book.

I will characterise what Taylor actually does in “Reconstructing Rawls” briefly as follows. In the book Taylor criticises Rawls for deviating from the Kantian foundations of *Justice as Fairness*; he proceeds to try and show that the foundations are indeed Kantian, and that his (Taylor's) subsequent reconstruction of the theory follows a line that Rawls could have taken, but did not. He then reshapes the theory into his own account of *Justice as Fairness*, based on the Kantian idea of moral autonomy – and even more fundamentally, upon Kant's practical postulate of freedom.

In the following discussion we need to keep in mind that it is Taylor theory, and his understanding of Rawls (and, Kant) that is primarily being discussed. I will try to avoid questions of whether or not Taylor's Rawlsian-Kantian theory is fair to Rawls as far as possible and only focus on the ideas put forward. Rawls' views have of course been presented and will be mentioned again; this is done to show what Taylor does and to be able to discuss the merits and drawbacks of his theory. It is however beyond the scope of this paper to make an in depth assessment of the merits and drawbacks of Rawls' theory in its different stages of development, so I will not attempt to do so.

In the following sections I will start by showing how Taylor reshapes Rawls' theory, focusing on how the regulative interests in the original position are amended and added to, and the

subsequent attempts to show that if we understand the contractors in the original position as being motivated by interests in moral and personal autonomy (as well as self-realisation) we have a much more stable base for the principles of justice and their internal priority.

I will criticise the construction on a few points, and overall I want to address the question of whether or not this is a viable way to defend the principles of justice, and the FEO in particular. I will at the onset say that I have serious doubts as to this, not because I think that this kind of Kantian constructivism doesn't hold promise, but because the process of trying to defend specific principles on grounds they weren't constructed on in the first place is a very backward, and *ad hoc*-seeming, process.

2.1 Highest and higher-order interests becomes moral and personal autonomy.

We are told in "*Reconstructing Rawls*" that understanding the contractors in the Original Position is understanding the Kantian conception of the person that underlies Rawls' theory. Taylor writes: "Most basically, the OP is intended to capture the Kantian idea that moral principles are "the objects of rational choice," that is, they are constructed rather than discovered or intuited, and they are therefore emblematic of our positive freedom or autonomy" (RR p. 28)

It is this autonomy that is the starting point of the contractors in the original position, and it is this conception of autonomy that Taylor will amend in order to (among other things) try to save the FEO principle. Taylor notes that Rawls assumes that all persons possess two moral powers, namely those of reasonableness and rationality. It is these moral powers that entitles them to equal justice - as Rawls says in *Theory* that moral personality is at least a sufficient condition for being entitled to equal justice. These are the constructivist foundation building blocks that are used in deriving the principles of justice from the Original Positions, so we need to look closely at how Taylor understands and then redefines them in his Kantian reconstruction. (RR p. 22)

a) Reasonableness – Moral Autonomy

Reasonableness is understood as the capacity to understand, apply and act from the principles of justice, I.e. it is *to be able to have a sense for justice*. Reasonableness is closely tied together with the idea of full autonomy, writes Taylor, which is the autonomy "achieved by individuals in a well-ordered society when they affirm, publicly recognise, and act from the principles of justice that would be chosen in a suitably characterised OP". And Taylor expands the first moral power of reasonableness to include the Kantian conception of self-authorship, moral autonomy, which he

says it's nearly identical to anyway. He goes on to say "if our authorial capacity were not part of our self-conceptions as moral persons, we would be incapable of imagining the development of it in a well ordered society" and so, he means to say, the fuller capacity of moral self-authorship must be contained within Rawls' conception of the person, and the first moral power of reasonableness. Reasonableness then, becomes moral autonomy in Taylor's reconstruction, and it is defined as "the choice of maxims of action consistent with and respectful of self-authored moral imperatives". (RR p. 22)

We can understand moral autonomy in Taylor's account to be closely tied to the original formulation of the categorical imperative: "act only according to that maxim through which you can at the same time will that it become a universal law" (GMM p. 34) a link that Rawls also points to in *Theory*, illustrated by the following quote:

"What distinguishes justice as fairness is how it characterises the initial situation, the setting in which the condition of unanimity appears. Since the original position can be given a Kantian interpretation, this conception of justice does indeed have affinities with idealism. Kant sought to give a philosophical foundation to Rousseau's idea of the general will. The theory of justice in turn tries to present a natural procedural rendering of Kant's conception of the kingdom of ends, and the notions of autonomy and the categorical imperative" (TJ p. 233)

In the course of examining Taylor's reconstructed *Justice as Fairness* it will be made quite clear that his account of what the interest in moral autonomy entails is the ultimate basis of the theory. Indeed, he does precisely what I referred to Rawls in section 1 as saying shouldn't be done: he derives the principles of justice from a necessary truth. The interest in moral autonomy is given because of Kant's practical postulate of freedom.

He renders the postulate as Kant's assertion that *we must presuppose the freedom of our wills on practical grounds*. Something about the way human agency works, it goes, makes it necessary for us to believe that our power of choice is undetermined by alien forces. That is, in the actual choice situation – when I am choosing, I "necessarily assume that my choice is not limited to one option, that my choosing is not a charade" and while I may think differently, and perhaps be a convinced determinist at other times: being able to make choices presupposes that I consider my choice a choice, and therefore my will to be free. It doesn't matter whether the will actually *is* free or not: all that matters is that in virtue of practical reason I must consider it to be free. (RR p. 303-305)

But how do we get from the (possibly) necessary conviction of free will to the interest in moral autonomy? Taylor describes the process roughly as follows. Kant claimed that causality is a category of understanding, along with quantity, quality, modality etc – and we can only experience things as having these categorical features: through them alone we apprehend something as an object. Cause and effect, then, is something that is universal, necessary and law-governed; “Categories like causality are part of the internal architecture of our theoretical or speculative reason; consequently, we necessarily seek the objectively lawful causality in our experiences – including observations of active human wills.” Because of this, the will must be governed by objective principles. **But, since we must consider our will to be free, these objective principles cannot be outside impositions** – but rather, the only alternative left if we must both view the will as free, and everything governed by causal objective principles is to view the will as setting its own principles. These would be the imperatives generated by practical reasoning. And these imperatives are the ones that Taylor reinterprets the regulative interests in the original position against. (RR p. 14-15)

To sum up: the interest in moral autonomy is derived from our interest in self-legislation, and our interest in self-legislation isn't really an interest, but a function of our practical reason – of the fact that we necessarily view our wills as free and therefore setting its own principles. In a hypothetical original contract situation then, where the contractors were to choose principles of justice to govern society, they would necessarily choose principles that would guarantee their interest in moral autonomy, because the way the Original Position works, with its Veil of Ignorance, is to ensure that the contractors don't choose heteronomously – but that their choices are products of their pure practical reason and not natural contingencies.

And we will remember that the way we arrive at the moral law according to Kant is by abstracting away from the personal differences among rational beings – indeed, the will can only be considered free when it isn't affected by heteronomous influences, or alien forces. (GMM 4:432)

b) Rationality – Personal Autonomy

Next comes the moral power of **rationality** which, as we will see, Taylor ties closer together with Kant's hypothetical imperatives. Taylor writes "Rawls defines the second moral power of rationality as the "capacity to form, to revise, and rationally to pursue a conception of the good"" (RR p. 24)

The second moral power of rationality consists in deliberative rationality and creative self-authorship, and is, Taylor writes, a "variation on the contemporary concept of personal autonomy". Deliberative rationality and creative self-authorship works together to create a plan of life; one that

is based on our interests and desires, but in a, so to speak, rational fashion. Taylor cites Rawls from *Theory* in describing the process of creating a rational plan of life, the second moral power: "The aim of deliberation is to find that plan which best organises our activities and influences the formation of our subsequent wants so that our aims and interests can be fruitfully combined into a scheme of conduct. Desires that tend to interfere with other ends, or which undermine the capacity for other activities, are weeded out; whereas those that are enjoyable in themselves and support other aims as well are encouraged" (TJ p. 360-61, RR p. 25)

Against this background Taylor reinterprets the highest-order interest in rationality as the interest in personal autonomy. Personal autonomy is in turn to be understood in light of the Kantian *imperfect duty of virtue to self and others*. Through this, personal autonomy can be said to be both ultimately autonomous (in that its origins are in the autonomous will) and substantially subjective. In Kant's account, we have imperfect duties both to our selves and to others, the basis of which is respect for humanity. Humanity is defined by the ability to set ends. This leads us to the *imperfect virtue to self of natural perfection*, and the *imperfect virtue to others of beneficence*. The duty to self of natural perfection impels us to develop our talents and skills – but, Taylor writes “we have great discretion with respect to which talents we should develop and how much we should develop them and [...] our occupational choice largely decides which set of skills we develop”. With regards to the imperfect duty of beneficence, “if the capacity to set ends defines humanity, then we must help others to achieve their (permissible) ends in order to respect them fully” though of course, here also we enjoy great discretion as to how we choose to do this – and so personal autonomy is substantially subjective while ultimately autonomous in origin. (RR p. 78)

Taken together, Taylor believes that the imperfect duties of natural perfection and beneficence shows how Kantian personal autonomy should be understood. It shows how we can fashion our plans of life in line with our particular “tastes, opinions, goals, values and preferences”, while simultaneously doing so in line with our duties to self and others, sprung from pure practical reason. (RR p. 79)

c) Derivative interests (?)

Apart from moral (reasonable and rational), Rawls also conceives of persons as free and equal. Taylor states that these two aspects of the person are "largely derived from the first one" (persons as moral). What it means to conceive of persons as free in this Rawls-Kantian way is described in three facets. (RR p. 26)

Firstly, people are entitled to make claims on institutions and society as a whole simply in

virtue of their humanity. Taylor ties this together with the third formulation of the categorical imperative (and this is what it means that the idea of people as free is derived from the idea of people as moral) regarding the treatment of humanity as an end and never only as a means. And humanity for Kant is said to be characterised, above all else, by the capacity to set ends, whatever they may be. The second aspect of freedom is said to be the "free revisability of our conception of the good and associated plan of life" and the third facet of freedom is (in a wonderfully Kantian fashion) responsibility; as long as free revisability holds, "citizens are capable of adjusting their aims and ambitions in the light of what they can reasonably expect and of restricting their claims in matters of justice to certain kinds of things" Taylor quotes Rawls as saying. (RR p. 26-7)

Equality, the "third and final component in Rawls' conception of persons" is tied together with the fourth formulation of the categorical imperative. From the first moral power of reasonableness, Rawls is said to ground equal worth, because "all [persons] view themselves as equally worthy of being represented in any procedure that is to determine the principles of justice that are to regulate the basic institutions of society". (RR p. 27)

d) Where is he going?

The cornerstone in Taylor's theory is clearly the reinterpretation of reasonableness and rationality as the interests in moral and personal autonomy. Of the two, the interest in moral autonomy is by far the most important one, because it is the interest in the development and exercise of this moral autonomy that ultimately determines what other interests are permissible. Taylor argues that his reconstruction of Rawls is not too far a departure from the original theory, but he is of the opinion that unless a more stringent Kantian conception of the person is maintained, it will not be possible to ground the lexical priority of the principles or even justify the Fair Equality of Opportunity principle.

Before moving on I would like to stop and make a reflection on what Taylor has actually done in translating reasonableness and rationality into moral and personal autonomy. Expressed simply, reinterpreting reasonableness as moral autonomy in the Kantian sense is quite a bold move – certainly it stands in stark contrast to Rawls' assertion that the principles of justice cannot be derived from necessary truths. The line of defence he has chosen is to try and show that the principles of justice are necessarily chosen the way they are as a result of our practical reason. This certainly brings the theory in line with Kant, but it does mean that much of what Rawls says about the interests that motivate the parties in the original position seem to become unnecessary, and more to the point, the role of the processes leading up to reflective equilibrium is much diminished. If we

can ground the principles of justice in necessary truths there is of course less reason to take our considered judgements into consideration – if they don't correspond to what can be derived from the necessary truth that is. This of course, is in line with Taylor's overall aim to take Rawls' theory away from the later-day Rawls' alleged relativism. The full scope of what this means and how Taylor wants to do this will hopefully be made clear by the end of this paper.

e) Adding Self-realisation

Reinterpreting reasonableness and rationality and dismissing the other regulative interests as mostly derivative isn't enough to justify the Fair Equality of Opportunity principle though. Were one an ungracious philosopher, one could perhaps take this to be an indication as to the feasibility of the principle in general – but let's be generous for the time being.

Taylor adds self-realisation as an element of personal autonomy, but a quite distinct element. It is, he says, a part of personal autonomy that has personal perfection as its subject. Self-realisation, he says "is in effect [personal autonomy] with a restricted range of application: instead of dealing with the wide array of desires, values, ends etc that can serve as the raw materials for self-authorship, [self-realisation] focuses on the development and possible perfection of our myriad skills and capacities, whether they are intellectual, physical, or spiritual". We should understand self-realisation, it seems, as a perfectionist self-development process, in which we typically struggle against internal barriers "including akrasia, myopia, and risk aversion" and is as much a part of "our autonomy as the overarching struggles for authenticity¹ and morality". (RR p. 84-5)

The reason self-realisation needs to be a distinct part of personal autonomy, we are told, is because of its object, which would be our skills and capacities, or in Rawls' vocabulary: natural primary goods. These are goods that aren't under the direct control of the basic structure of society, even though they may be influenced by it. They are influenced by the social structures, and cultivated according to our rational plan of life, but they are, importantly, internal in nature. "Whereas social primary goods are possessions, natural primary goods are best regarded as part of our identity - they are something we are rather than something we have" (RR p. 85-6)

Perfectionism is in this fashion made a subordinate part of personal autonomy, which means that the end of personal perfection isn't discretionary. While we are able to choose how we are to cultivate our skills and abilities (and which ones of our possible skills to focus on etc) the end of self-perfection is dictated by moral autonomy. Moral autonomy guides personal autonomy, of which

¹ Taylor sometimes speaks of "authenticity" in relation to the formation of a (rational) plan of life. As far as I can tell it doesn't bring anything new to the table, but seems to mean simply authentic as autonomous in origin.

self-realisation is a component, and therefore it is in line with and contributes to the overall interest in moral autonomy. (RR p. 85)

Rawls explicitly rejects perfectionism in *Theory*, defining it as either the principle that society should be arranged so as to maximize achievement of human excellence in art, science and culture, which is seen as an overriding value, or the more moderate version where perfectionism can be weighed against other standards. Taylor's description of perfectionism seems to avoid the objections that Rawls has against perfectionism in *Theory* since the main problem Rawls seems to have with perfectionism is that it puts some ideal above or beyond justice. Taylor's perfectionism is valuable *because* it contributes to the overall demands of justice – because it contributes to moral autonomy. If it didn't, we must conclude, it wouldn't be considered valuable. (TJ p. 285-289)

Whether this is taking the differences between individuals seriously is, I suppose, up for debate. That debate would, however, be a bigger discussion concerning how constraining a Kantian conception of the person really is, and time and space doesn't permit me to have that discussion here.

2.3 The Principles of Justice and their priority

Before we get into the principles of justice it must be noted that there is a meta principle that provides the framework of the principles of justice. This is the principle of the Priority of the Right over the Good. For Rawls, ""The concept of the right is prior to that of the good," regardless of the goodness under consideration". Here, I believe, we can see one of the fundamental points of conflict with Arneson and Alexander, for as Tyler asks "Why can't justice be sacrificed for the sake of welfare or perfection or piety, at least when the exchange rate is highly favourable?". In Taylor's account this is because of Rawls' Kantian constructivism and his conception of persons as morally autonomous, something that Rawls isn't stringent and clear enough about according to Taylor, and so we can understand why the question would arise. But Arneson and Alexander could also ask why persons couldn't equally be conceived of as a "sensors being, an artistic creator or a child of God" and why these conceptions shouldn't take priority in formulating the principles – why is the person as an autonomous moral legislator the given starting point? The answer is hopefully given in my description of how Taylor derives the interest in moral autonomy from Kant's practical postulate of freedom – though it is of course not uncontroversial. (RR p. 42-3)

- a) The interests in moral and personal autonomy yields the liberty principle

The principles of justice is chosen to guarantee the regulative interests, and the internal priority of the principles of justice reflects the weight of the interests. The connection between the first principle of justice – the priority of liberty – and Taylor's reconstructed interest in reasonableness, the interest in moral autonomy, is easily understood I think. The parallel is expressed clearly by Taylor: “Just as moral autonomy can be understood on Kantian grounds as the legislation of morality for a kingdom of ends, so in the context of right it can be understood as participation in republican self-government, whether directly (as legislator or bureaucrat crafting laws or regulations, respectively) or indirectly (as a voter judging and choosing between legislative programs of parties and candidates)”. (RR p. 133)

Political liberty should not be viewed as means to ends. Taylor explicitly rejects Rawls' attempts to justify the priority of liberty on instrumental grounds; on the grounds that political liberties are necessary for the protection of civil liberties, which in turn would be necessary for the advancement of ones conception of the good. The general argument against this is that the priority of liberty cannot be sustained by this line of argument, because “even if the interest to be protected by political liberties (e.g. civil liberties) has the requisite lexical priority, the priority of political liberties will not be justified by implication unless they are strictly necessary for such protection. If they are merely strongly contributory towards such protection, then other means of protection can potentially be substituted for them”. We might imagine for instance that instead of spending the money and resources needed to protect the fair value of ones political liberties, it would be cheaper and easier to try and guarantee a just legislative process by investing in publicity and transparency, so that unjust laws would be “likelier to be identified and resisted” and if this was indeed a less expensive way to go about guaranteeing just legislation, then the lexical priority of political liberty couldn't be sustained, Taylor argues. (RR p. 127-28)

Indeed, Taylor's has to contradict Rawls on quite a few points to defend his conception of the relationship between the regulative interest in moral autonomy and the principle of liberty. More notably even than the rejections of Rawls' attempts to defend the priority mentioned above is the rejection of Rawls' assertion that the the two powers of reasonableness and rationality are not ranked. Taylor recognizes that his theory is only valid if the capacity for a sense of justice is hierarchically ordered above the capacity for a conception of the good. (RR p. 135-36)

To briefly sum up: in Taylor's account the principle of equal liberty is chosen and lexically ordered by the contractors in the Original Position because it reflects and in a way *constitutes* their interest in moral autonomy, in self-legislation.

The interest in personal autonomy is reflected both in the liberty principle and in both of the second

principles parts (The FEO and the DP). In the case of the liberty principle we note that Taylor construes it as being primarily concerned with political liberties, and that these are (contrary to what Rawls says) to be hierarchically ordered above civil liberties. Civil liberties are of a couple of different varieties it seems: either they are necessary to secure political liberties, or they are not. They are, however, always to be understood as having priority over the FEO and the DP – regardless of whether the civil liberties in question are of the stronger or weaker kind. (RR p. 135-6)

I will talk more about the securing of the fair value of the political liberties later; for now I will only note that my understanding of the difference Taylor makes between the two categories of civil liberties are based on whether they are strictly necessary for guaranteeing the fair value of the political liberties or not.

Generally we are to understand civil liberties as tied to the exercise of rationality, i.e. personal autonomy: “civil freedoms of speech and assembly, liberty of conscience, and freedom of thought are *directly* supportive of the creation and revision of plans of life” we are told. (RR p. 166)

In Taylor's account then, the first principle of justice, the equal liberty principle, is chosen by contractors in the original position in order to secure their highest order interest in moral autonomy, as well as give the basis needed for the exercise of personal autonomy. They recognise however, that there can never be a trade-off of political liberties (guaranteeing moral autonomy) for the sake of civil liberties (guaranteeing personal autonomy) – the trade-off can only go the other way. We are left wondering how we are to handle trade-off-situations cases concerning civil liberties that are necessary for the exercise of political liberties. We might also ask ourselves if it wouldn't be more practical to cut the first principle in two, since it apparently deals with two very distinct sorts of “liberties”, grounded in two different interests. And add to this the fact that the FEO principle is introduced because of our interest in self-realisation, which itself is grounded in the interest in personal autonomy, it seems as if Taylor makes his own account unnecessary confused in order to adhere more closely to that of Rawls.

b) The interest in self-realization yields the Fair Equality of Opportunity principle

The defence of the Fair Equality of Opportunity principle in “Reconstructing Rawls” seems to suffer from essentially being a reworking of the 2005 article mentioned above. The problem with that is that we find ourselves standing in a argumentative limbo between a defence based on Rawls' actual writings and a defence based on the strict Kantian reconstruction in the rest of the book.

While it might be interesting to speculate and perhaps to extrapolate from Rawls' writings about the Aristotelian principle and the Humboldtian concept of social union, the main defence

doesn't really lie here any more.² Now that Taylor have made his reconstruction of the foundation, that is the path we must follow. So, the question becomes how the interest in self-realisation, as a distinct part of the interest in personal autonomy yields the FEO principle – in the form it is rendered by Rawls no less.

It seems that the arguments for self-realisation as a distinct part of personal autonomy is just those that was described above. It is, we are told, because of the special objects it concerns itself with that self-realisation needs to be distinct. As far as I can see this argument isn't developed much further, except perhaps to say that it (the interest in self-realisation) is “tightly connected with the lowest rung of Kant's ladder of necessitation – rules of skill – in his own model of finite rational agency”. Just pointing to the old Prussian hardly constitutes an argument though. When connecting moral autonomy and personal autonomy to the categorical imperative and the concept of hypothetical imperatives, that served the function of either illustrating what was meant or borrowing justifications from Kant. Perhaps Taylor means to use Kant's argument for the rules of skill to defend why self-realisation is a distinct part of personal autonomy, but if so we are left waiting for that argument. (RR p. 85)

We must conclude that what Taylor means to say however, is that contractors in the original position would choose a principle intent to protect their interest in personal perfection, an interest that is ultimately derived and guided by the interest in moral autonomy (by the duty to self and the duty of beneficence). Furthermore, Taylor obviously means to say that the principle that is to guarantee the interest in personal autonomy (the liberty principle) isn't enough to guarantee the interest in self-realisation. Personal perfection, it seems, needs a special principle, albeit subject to the liberty principle.

Only from this point of understanding can the arguments in the chapter dealing with the FEO principle (based on the article) be understood I believe. In it we are mostly faced with the question of why on earth the interest in self-realisation, personal perfection, would be institutionalised by the contractors in the original position in the form of a principle that focused on guaranteeing that (job) careers were open to talents and that chances for positions were made substantially equal. The focus on jobs must be defended, and that defence (found in chapter 5) can be summed up roughly as follows.

Taylor explains the importance of the FEO by stating that the positions associated with those social unions that are part of the basic social structure which he identifies as “governments, private and public corporations, universities, NGOs, and so forth has such a large role in peoples lives. The

2 For a variant of such an interpretation I recommend for instance Eyal, Nir “*Perhaps the most important primary good*”: *self-respect and Rawls's principles of justice*” *Politics Philosophy Economics* 2005 4: 195

social duties associated with them are of great importance, and of course this is the primary arena for the development of skills and the pursuit of virtuosity. The connection between the Aristotelian principle and the FEO is said to be just that: this is the primary arena for the pursuit of virtuosity. (RR p. 180-81)

Given that Taylor describes the organisations that belong to the basic social structure as he does, he must then still face the critique raised by Arneson and Alexander. Of course, I think that it is reasonable to have a wider definition of what is to be included in the basic social structure than Rawls (perhaps) has, as I mentioned in section 1.2. Defining the interest in self-realisation as he does, as something given us by pure practical reason and being derived from the universal law of the categorical imperative: well, it makes sense that it should produce a principle that hits a bigger target.

We can also note the reference to Kant's apparent conviction that “the employment context [is] the primary arena for self-realization, the institutional space that allows us to develop our myriad intellectual and physical skills for public purposes” (RR p. 189)

Indeed, it might be that the employment area fits the imperfect duties of self development/perfection and beneficence especially well, considering that we can at the same arena develop our skills and put them to good use. Providing, of course, that the work we do really is a good one; Imagine that I was a lawyer and extremely skilled at getting despicable criminals of the hook in various contexts. I would certainly be developing my skills, though perhaps not for 'public purposes' - but that is a difficult evaluation: when does the development of our skills within the framework of our job work in line with both the duty of self development and the duty of beneficence? Perhaps the question is too simplistic – but even still I am left wondering how restricted our choices would be if they must be “guided” ultimately by the interest in moral autonomy.

Regardless of this question of working for good and bad, the overriding question is still whether or not we should accept the special role assigned to work/employment by the FEO principle. This, I believe, can be answered by reviewing the following:

What would be needed for the contractors in the original position to choose the FEO principle would be that a) they actually had the interest in self-realisation such that Taylor describes it (of course) and b) that the subject of the FEO principle, that is; offices and positions in the basic social structure (in governments, private and public corporations, universities, NGOs, and so forth), are *necessary* to guarantee the protection of that interest.

I am of the opinion that b) is true. Given the fact that a) is true, that persons have the interest in self-perfection the way Taylor conceives of it, I think the contractors in the original position

would choose a principle guaranteeing the primary arena for such activities. Of course the whole thing is quite backwards as it stands: Taylor has the finished principle formulated by Rawls and then builds a suitable interest to justify it. I will call this my *ad hoc-charge* against Taylor.³

Had he started with the interest I think the principle would have been formulated differently. However I do think that it would have included offices and positions roughly as described by the FEO principle - but it would perhaps be expanded to include other areas as well: the workplace is after all not the only arena where self-perfection is developed, although it is perhaps the largest and most important. We could say that there are serious doubts as to whether c) a principle of justice formulated as the FEO is, regulating offices and positions in the basic social structure (in governments, private and public corporations, universities, NGOs, and so forth) is *sufficient* to guarantee the interest in self-realisation, obtains. Is it the best principle to do the job?

This all hinges on whether or not the contractors in the original position is aware of what role the workplace has in human society, if this is included in the general facts about the world that they are allowed behind the veil of ignorance. Would the contractors know this? Perhaps it is reasonable to say that they would. If they are aware of human nature, understood (among other things) as the knowledge about the necessary interests derived from pure reason, then perhaps they would have to be aware of the arenas where those interests can be developed.

Before moving on I would like briefly to mention an objection that I am thankful to have had Simon Rosenqvist point out to me. Might it not be conceivable, the objection goes, that in some cases, fair access will hinder the promotion of excellence, since some people will end up in worse places (from a perfectionist standpoint) if they have equal access, rather than a restricted access, to positions. If we want development of excellences, we should want a society which puts everyone in their right place, not one where people have equal access to different positions.

I believe that this objection is deflected when we note that the perfectionism in question gains its legitimacy in virtue of its contribution to autonomy: in this sense we could say that increasing perfectionism at the cost of autonomy would constitute a trade-off from a lower interest to a higher (trading autonomy for self-realisation) which isn't allowed, and wouldn't make sense if self-perfection is valuable *because* it contributes to autonomy.

³ Again I thank Eric Brandstedt for suggesting this wording.

Part 3

3 Problems and advantages of Taylor's rendition of the theory

In this chapter I will address some of the specific critiques raised at the onset of this paper and discuss the advantages and problems with the possible answers that Taylor's theory gives or can potentially give. The first question that I wish to address is that of the cost of the FEO principle, in a sense this is the basis of the utility argument. Put simply: *will the cost of having the FEO principle be lexically prior to the difference principle become too great, even against the background of the rigidly Kantian construction that Tyler proposes?*

a) The black hole-argument – is it a red herring?

One of the main charges that both Arneson and Alexander raises is that the exchange rate in many cases would be too high; that prioritising the FEO principle over the Difference Principle would require society to funnel so much resources into securing Fair Equality of Opportunity that the ability to better the situation for the worse off via the difference principle would be seriously impeded.

These kinds of “black hole”-arguments are said by Taylor to be red herrings. He is evidently of the opinion that much of the critique that Arneson raises can be fended off rather easily by pointing out that Rawls' theory isn't a maximising theory. If it was, then actually, the big problem wouldn't be the priority of the FEO, but the priority of the liberty principle, because if the higher principle always won out in the fight over resources, very little would be left over for the others at all. But since Rawls doesn't conceive of the right consisting in maximising the good, we avoid this. “There is no requirement to maximise our capacity for moral autonomy - or our capacities for personal autonomy and self-realization” since, as Rawls says “nothing beyond the essential minimum is required”(RR p. 145, TJ p. 442)

However, this statement by Rawls (“nothing beyond the essential minimum is required”) is not made in the context that Taylor implies. Rawls is talking about who are entitled to equal justice, and not about the distribution of goods in that particular context. More importantly – Taylor's theory isn't Rawls', and he himself hasn't provided an explanation for why his rendition of the theory wouldn't require maximisation of the goods connected to the regulative interests.

Moral autonomy is the guiding interest of all other interests in Taylor's account of the theory, and the extent to which other interests are to be protected is determined by whether or not they are

in line with the interest of moral autonomy. On the surface of it, it would seem a great difficulty for Taylor to explain why we aren't obliged to maximise moral autonomy at the expense of fulfilment of other interests.

In Rawls's original theory interests are balanced against each other, and the strict hierarchy that Taylor proposes is notably absent (for example in the case of political and civil liberties mentioned above). Because of this I think that Taylor owes us an explanation as to why maximisation of those goods connected to moral autonomy isn't needed.

The goal of Taylor's theory is surely to create a theory of a social framework that facilitates moral autonomy (since it is the necessary willing of moral autonomy that motivates the creation of the framework theory in the first place) so to say that this (moral autonomy) doesn't necessarily need to be maximised simply sounds strange as an argument in this context.

If the argument is, as it seems to be, that the “black hole-argument” is a red herring because actually, we are only required to arrange things so that there is a minimum of moral and personal autonomy as well as self-realisation, what does this even mean? It would be absurd to say that in the case of the highest-order interest in moral autonomy, and the corresponding priority of liberty, we only need to assure a bare minimum of legislative influence. What we need to do is obviously to ensure exactly the right amount of political influence so as to fulfil the requirement of the liberty principle, thus guaranteeing the interest of moral autonomy. Of course, we can only provide a framework, and we aren't able to (or I suppose, *allowed*) to make sure that everyone in society *actually* satisfies their highest-order interest in moral autonomy and take part in the legislative process (vote, join a party or whatever) but even so, the requirements on the framework must reasonably be that everyone should *be able to*. This “problem” returns below when we consider the requirements of the liberty principle in Taylor's account.

I am left to wonder why the institutional framework that supports the interests in personal autonomy and self-realisation should be any different. With this in mind I have to conclude that it is at least not obvious that the black hole-argument is a red herring for Taylor's theory. I think it has to be defended against, and it may be possible to do so by showing that this type of critique isn't really that damning. The way to go about it could be to show what is required in a non-ideal society as opposed to an ideal.

b) The second answer: The threshold

There is a more successful defence against Arneson's charges, or at least a way to make them not seem so damning, that Taylor can and does use, just as Rawls originally does. The defence in

question is of course Rawls' assertion that the priority of the basic liberties only hold under "reasonable favourable conditions" that is to say, they may be restricted if it is necessary to do so in order to get to a level of development where the principles can hold. There are of course extensive difficulties in deciding when these criteria have been met, but I propose we could say that *it cannot be the case that due to the lexical priority of the principles, resources can be drained from the fulfilment of basic needs in such a way that the basis for maintaining a framework where the principles of justice holds are eroded.*

When Taylor deals with the charges against the FEO of creating "black hole" drains on resources directly he writes: "As I have argued [...], the three moral powers of reasonableness, rationality, and self-realization must only be developed "to a certain minimum degree... nothing beyond the essential minimum is required" [...] There is certainly no demand to maximize their development, whatever that might look like, so there is consequently no threat of all economic resources being consumed in an insatiable quest for maximal self-realization. Finally [...] insofar as basic consumption needs (e.g. food, shelter, clothing) are preconditions of the *effective* development and exercise of our capacity for self realization, the lexical priority of FEO must give way to them."(RR p. 183)

It is as I said above, dubious as to whether the first of these two arguments works within the rigidly Kantian framework Taylor proposes, at least with regards to the first principle, though the second one seems to have more weight. Obviously basic needs have to be met in order for it to be possible for people to satisfy interests of all three aspects of autonomy. That still leaves open the question of how well off we need to ensure that the worst off are, or differently put, when the threshold level is passed and the difference principle give way to the FEO principle (assuming of course that application of the difference principle is the most efficient way to ensure a certain level of welfare for people in order to be able to pursue the interests).

Taylor writes that Rawls must have something similar in mind for the application of the FEO principle as in the case of the threshold condition for the basic liberties, which would be an "unspecified level of social, legal and economic development" so that "adequate sociopolitical and material resources must be available before the priority of FEO goes into effect, where "adequate" means whatever level is necessary for the liberties associated with FEO to be "effectively exercised""(RR p. 185)

Taylor clearly mean that resources should, or must, be devoted to the principles and the interests which they protect in the right order. He writes that "Once the requisite level of social, legal, and economic progress has been reached, however, the priority of liberty comes into play; given the lexical priority of EL over FEO, an increasing social resource base must first be used to

secure the priority of liberty. As the resource base grows though, a point is eventually reached where the priority of FEO comes into effect"(RR p. 186)

This pretty much leaves us back with the first problem I discussed though. Since it is hard to see how moral autonomy (and the corresponding priority of Equal Liberty) could be required to be developed only to a minimum level, we are left with a suspicion that if resources are to be allocated to the principles in order of importance, that might indeed turn them into "black holes" for resources.

We can imagine a fictional early-modern society to illustrate, roughly corresponding to Sweden in the nineteenth century . We could indeed justify the lack of equal liberties (equal voting rights and the securing of a properly democratic process) by saying that we need to restrict these things until we have reached a level of economic and social development where this can be afforded (sacrificing liberty for stability as it were). But let us assume that we (perhaps we are the enlightened king in this case) decide that that level has been reached. Then the priority of liberty comes into play, we recognise. So we set about to secure Equal Liberty. This is of course a daunting task, since we need to make sure not only that there is formal equality (that everyone has equal right to vote and equal rights to offices connected to the exercise of political power and so on) we also have to make sure that the political liberties have their "fair value" guaranteed. We need to make sure that "people have a genuinely equal chance to "determine the outcome of...the constitutional process that establishes the laws". The interest in moral autonomy can demand nothing less. (RR p. 122)

This will obviously be extremely costly. Taylor writes that the incorporation of a specialised FEO principle within the Equal Liberty principle (that this securing of the fair value of equal liberty consists in) will require (among other things) public financing of political parties, public support for political debate, and **"state efforts to prevent the concentration of property and wealth, all of which are needed to preserve a level political playing field and prevent the democratic process from being hijacked by powerful special interests"**(RR p. 122, my emphasis)

We will notice (in our hypothetical situation as the Enlightened monarch of the Swedes, Goths and the Wends) that indeed, substantial meaningful Equal Liberty requires even more than this though. If it is, as Taylor writes that "Political Liberty [...] is not just a means to the development and utilisation of our capacity for moral autonomy, but in some ways *is* this capacity in political institutional form" then we can spare no expense to make sure that everyone has de facto equal chance of exercising moral autonomy, now can we? (RR p. 133)

Indeed, we find ourselves thinking that we need to construct quite a radically egalitarian society, for surely, no one can be expected to be allowed to keep more than is necessary of the fruits

of their labour, when these resources could be used to ensure someone else's education (and improvement of life situation in general) in order for this person to be able to take part in the democratic process in a fashion that properly expresses their interest in moral autonomy? So we find ourselves in our Royal Person wondering when we became such a socialist, and also, whether the other principles within *Justice as Fairness* in its reconstructed form are even necessary, now that we have realised how extensive the first principles demands really are.

Stepping out of this example and returning to the discussion of what threshold conditions that needs to be fulfilled before the FEO principle is to take priority, I find that the answer really seems to be something along the lines of what I suggested above: *that it cannot be the case that due to the lexical priority of the principles, resources can be drained from the fulfilment of basic needs in such a way that the basis for maintaining a framework where the principles of justice holds are eroded.* But beyond that, it really doesn't seem to be permissible to allocate resources in order to satisfy needs that rank below the interests regulated by the principles.

It seems to me however, that even if we say this it might still be that with this strictly Kantian interpretation of the basis for the principles, they would become demanding in such a way that the effects would be ones that the consequentialists could hardly think bad, even though they might disagree with the principal foundations. Let me explain further what I mean by this.

When writing about what substantive FEO requires, "namely, that all family and class inequalities be neutralised" Taylor indicates that in order to achieve a level playing field, the obligations of society will be quite substantial. He notes that "a disadvantaged family and class background, for example, if left uncorrected, may cause agents to rule out particular occupations (e.g. white-collar jobs) and related forms of training (e.g. a college education) from the start". Just as an apartheid state would impose heteronomous restraints on self-development by direct actions, so the state because of inaction could be responsible for heteronomous restrictions based on social contingencies in occupational planing, he writes. (RR p. 183)

This, to me, indicated that when the FEO principle comes into play, what it will mandate is quite far-reaching social policies, designed to secure substantial equality of opportunity. Where I to complain that this would lead to there being too few resources left for the difference principle to allocate to the worst off, I (if I wanted to defend Taylor's theory) could simply answer the following:

The worst off are guaranteed a minimum of resources in order to be able to pursue their interests of moral and personal autonomy, and indeed, self-realisation. Furthermore, the Equal Liberty principle and the Fair Equality of Opportunity principle ensures such a level of substantial equality for people to be able to both participate fully and in the democratic society and be able to

develop their skills that I have a hard time of seeing where the Worst Off could be lacking. We are required to spare no means to ensure that everyone can participate fully in the democratic processes of society, and since this reasonably means working to eliminate all kinds of inequalities and class barriers that stratify and alienate people, what more could be asked? And as if this wasn't enough we are required to do even more in order to ensure that people are equally able to develop their talents. Even if there is someone who might be said to 'fall behind' (perhaps the severely untalented as the critics suggest) she would not be allowed to fall so far back, as it were, so as not to be able to participate in the democratic society (so she would still need to be educated reasonably well, guaranteed a reasonable minimum standard of living and guaranteed time to participate in the democratic culture) and so you would have to be a pretty hardcore egalitarian to protest the situation.

One of the conclusions I draw then is that Taylor does indeed circumvent a lot of the critique that are being raised by Arneson and Alexander, but he doesn't do it in the way he himself intends to do it. On the surface of it it seems clear that what he wants and sets out to do is to show that there is a regulative interest, drawn from a (in his view) superior conception of persons – self-realisation – that motivates some of the costs that Arneson and Alexander argue are so unsavoury. And crucially, the way he actually does circumvent this critique begs a whole lot of other questions that may be much more troublesome.

3.1 Some specific objections

Now lets briefly look at what we might be able to say in response to some of the specific objections that were raised by Alexander and Arneson.

aa) The Alaska Case

I think the answer to this critique should simply be that were we to allow for the principle to be this far-reaching, that would create a massive resource-drain that would erode the basis for ensuring the liberty principle. This is perhaps an inadequate answer, but since the answer that Taylor seems to want to give (that maximisation isn't necessary) clearly fails to make sense in light of his reconstruction I am forced to extrapolate and suggest this instead.

bb) The Socialised Ambition Case

In this case it seems that Alexander and Taylor simply has differing opinions of what the FEO principle demands. Taylor writes, in connection to the demands by the FEO that social inequalities be corrected “a disadvantaged family and class background, for example, if left uncorrected, may cause agents to rule out particular occupations (e.g. white-collar jobs) [...] whereas in the apartheid case the narrow conduits along which victim self-development was directed were both designed and reinforced by the state, in this case the conduits are maintained by state *inaction*, a passivity that allows the heteronomous influence of social contingency to constrain occupational planning”(RR p. 183)

cc) The Unequal Abilities Case

In answer to this critique we could say a couple of things. First of all, we could say that talents are talents, and that we are to choose what skills and talents to hone and perfect in accordance with a rational plan of life, as described above. In Taylor's account, the FEO isn't really about *equalising* abilities at all, but providing what is needed for self-perfection to be possible. The extent to which different kinds of specific abilities is developed in every specific person will presumably be relative – primarily in relation to that person, and only secondarily in relation to the outside world (obviously the world's limitations and rules determine what kinds of skills that will constitute self-realising ones if developed, but that is the extent to which the outside world determines).

dd) The Wasted Resources Case

This objection have been dealt with above in several ways. But we could again consider the main point of the principal argument: “*If, however, equality of opportunity does not translate into better jobs or income for the worst off, it is difficult to understand why Rawls' rational contractors in the original position would value it*”. (Alexander, 202)

Given the fact that the intuitive basis for *Justice as Fairness* is said to be exactly that any inequalities should be able to be justified to the worst of in virtue of them (the inequalities) actually leaving the worst off better off than they would otherwise be - then Alexander's critique is potentially very potent. (TJ p. 13)

To me his critique very much seems to hinge on the general facts that the contractors in the original position is said to be aware of. They are not aware of who they are, or anything about themselves, but they are aware of general facts about the world and human nature and so on – to be able to choose principles at all. These facts, it seems to me, could be facts about human psychology

that would confirm that guaranteeing fair opportunity is necessary to facilitate the power of human ambition, and thus guarantee efficiency and the subsequent benefits to the worst off. The facts could just as easily deny this though. I don't know, I am not a psychologist, but I'll make the (weak) assumption that consensus has not been reached in the scientific community dealing with these matters yet.

Assuming that it can be shown that the FEO principle yields inequalities that can't be justified to the worst off, then it does seem that the only viable way of saving (or making sense of) the principle is to show that it is tied to a specific interest, on par with the interest expressed and protected by the first principle of fair liberty. In Taylor's account the answer to the question of why the contractors would choose the FEO principle even if it sometimes would yield results that will leave the worst off worse off than they could have been is simply: 'because of practical reason'. The intuitive basis for Taylor's *Justice as Fairness* could then be said to be quite different from that of Rawls'.

Conclusion

In this paper I have examined Robert S. Taylor's reconstruction of John Rawls' *Justice as Fairness* and in particular focused on whether or not his reconstructed theory can provide a strong defence of the Fair Equality of Opportunity principle. I have held that his grounding of the FEO principle in the interest in self-realisation could partly motivate the FEO principle, but that had he constructed the principles of justice after he had made the regulative interests in the original position clear, I believe the scope of the FEO principle would be somewhat different. I raised essentially the same critique (though less extensive) against the Fair Liberty principle, which seems ill fitted to handle the clear separation of political and civic liberties that Taylor insists upon. There is, I suppose, some value in trying to show how *Justice as Fairness* could have been, had the foundations been more boldly Kantian, but I do hope that Taylor develops the theory on his own and allow Rawls to be more of an influence rather than a template which he must fit principles into.

