This is the published version of a paper published in *Philosophical Papers*.

Citation for the original published paper (version of record):

Lindblom, L. (2018)
In defense of rawlsian fair equality of opportunity
*Philosophical Papers, 47*(2): 235-263
https://doi.org/10.1080/05568641.2018.1445550

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-148123
In Defense of Rawlsian Fair Equality of Opportunity

Lars Lindblom

To cite this article: Lars Lindblom (2018) In Defense of Rawlsian Fair Equality of Opportunity, Philosophical Papers, 47:2, 235-263, DOI: 10.1080/05568641.2018.1445550

To link to this article: https://doi.org/10.1080/05568641.2018.1445550

© 2018 The Author(s). Co-published by Unisa Press and Informa UK Limited, trading as Taylor & Francis Group

Published online: 29 Jun 2018.

Article views: 462

View Crossmark data
In Defense of Rawlsian Fair Equality of Opportunity

Lars Lindblom

Abstract: Richard Arneson argues that Fair Equality of Opportunity (FEO) should be rejected, since it is not only too weak and too strong, but also problematically meritocratic. The paper aims to defend FEO, and argues that it is not too weak, since, pace Arneson, it does apply to the problem of stunted ambition. The argument from meritocracy is shown to be based on a conflation of different senses of meritocracy. Finally, it is shown that FEO, correctly interpreted, gives intuitive answers to the examples put forward to bolster the too strong charge. It is concluded that Arneson’s refutation of FEO fails.

1. Introduction

This is a paper on the principle of Fair Equality of Opportunity (FEO). This principle plays a very important role for Rawls’s theory of justice—it is lexically ordered above the difference principle—but not much work has been done on it. For instance, neither of the Companion volumes that have been published about Rawls have any articles devoted to FEO (Freeman 2003, Mandle & Reidy 2001). One reason for this could be that Rawls himself says very little about the principle, and some of what he says is not easy to interpret. A second reason, however, I would argue, is the work of the one political philosopher who has given this principle the attention that it deserves. In a series of papers published over the last 15 years, Richard Arneson has investigated this principle from two directions. On the one hand, he has looked into the substantive content of the principle and argued that it should be rejected. On the other, he has investigated whether it would be chosen in the original position, and claimed that it would not (Arneson 1999, 2006, 2013a; cf. Arneson 2013b, 2015).
In this paper, I will discuss Arneson’s arguments concerning the substantive content of the principle and leave the issue of whether it would be chosen in the original position to another time. These arguments were first presented, and are most fully articulated, in ‘Against Rawlsian Fair Equality of Opportunity’ (Arneson 1999) and my focus will be on this paper, and the arguments presented therein. Arneson presents three rather overwhelming charges:

1. Rawls’s Fair Equality of Opportunity Principle should be rejected because it is too weak, insofar as it fails to condemn wrongful discrimination that operates through ambition affecting socialization.
2. Rawls’s Fair Equality Principle is also too strong, in that it condemns discrimination that should be deemed innocent rather than wrongful.
3. Fair Equality should be rejected because it represents a compromise with meritocracy, and wrongly constrains the wholehearted pursuit of that part of justice that seeks to maximize a function of the opportunities and liberties available to the members of society that gives priority to improvements for the worse off. (Arneson 1999: 109)

I shall argue that these criticisms rest on a questionable reading of Rawls, and that Arneson’s argumentative strategy has deep flaws. The aim of this paper, then, is to defend FEO against Arneson’s three charges. This aim also explains the outline of the paper. In Section 2, I present my reading of FEO, which positions the ideal of free and equal citizens at center stage. In the third section, I attempt to show that the Too Weak charge rests on a misinterpretation of the scope of FEO. For reasons of disposition I reverse the order of Arneson’s charges and turn in the fourth section to the Meritocracy charge. I will argue that it, first, rests on the conflation of the two very different ideas that (1) positions should go to the most suitable candidate and (2) the more talented ought to be rewarded for their talents; and second, that the idea that priority should be given to achieving FEO over the difference principle can be given a plausible reading, which can be used to defuse Arneson’s charge. In the fifth section, I discuss a number of
Arneson’s examples that aim show that FEO is Too Strong since it finds unfairness in areas where discrimination is best considered innocent. I claim that on closer inspection FEO gives the right answers, or at least as plausible answers as the principles that Arneson favors. In the final section, I sum up the results.

2. Fair Equality of Opportunity

The core normative content of John Rawls’s theory of justice—justice as fairness—is presented in the form of two lexicographically ordered principles. The first principle, in its final formulation, demands a fully adequate and equal scheme of rights and liberties, and says that when it comes to the political rights and liberties the fair equal value of these should be guaranteed, whereas the other rights are underwritten by the second principle. This principle says that:

Social and economic inequalities are to satisfy two conditions: first, they are to be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they are to be to the greatest benefit of the least advantaged members of society. (Rawls 2001: 42–43)

The first part, FEO, is lexically prior to the second part, the difference principle, but other than that Rawls is rather cryptic when it comes to the meaning of this principle. A reader might be excused for thinking that the only help he or she gets is the following:

[T]hose with the same level of talent and ability and the same willingness to use those gifts should have the same prospects of success regardless of their social class of origin, the class into which they were born and develop until the age of reason. (Rawls 2001: 44. Cf. Rawls 1999: 63; 1996: 5–7)

The obvious interpretation of this quotation seems to be that education (and all other institutions of the basic structure in different ways and to varying degree) should ensure that citizens at the same level of talent and willingness should have equal opportunities, but that citizens at different levels could have unequal sets of opportunities. This would make talent, ability and willingness the only aspects that justice allows to affect
individuals’ chances of attaining their goals. In consequence, sexual orientation, ethnic background, gender and class should not affect individuals’ prospects of success in realizing their opportunities. While this interpretation is clearly salient, it is not quite right; it does not make full sense of Rawls’s vision of democratic equality.

Samuel Freeman captures the egalitarian idea behind FEO in the following manner: ‘Being in a position to develop one’s capacities and talents, whatever they may be, is needed to maintain one’s status and self-respect as a free and equal citizen capable of social cooperation over a complete life.’ (Freeman 2007: 95) Freeman identifies the three major arguments for FEO in Rawls’s works as (1) The Aristotelian Principle, (2) the fact that it is needed to make sense of the difference principle, and (3) the ideal of free and equal citizens. That efficiency is not found among these reasons may be surprising. Obviously, there is a close connection between that value and equality of opportunity—if each person has equal opportunity, then the most suitable candidate will be hired for each position, which in turn will improve efficiency—but this is just a lucky contingency.¹

I shall not discuss the Aristotelian Principle (which says roughly that people prefer to exercise their developed faculties, and enjoy more complex tasks to less intricate undertakings) (Rawls 1999: 374–375. Cf. Taylor 2004). Using this principle in the argument for FEO might amount to a violation of the strictures of the political liberalism of the later Rawls, and, moreover, the arguments I will make stand on their own without the support of the Aristotelian Principle.

Let us, then, focus on the connection between FEO and the ideal of free and equal citizens. Rawls says the following:

¹ On Arneson’s interpretation, efficiency is major reason in favor of FEO. He argues against FEO that the value of efficiency has some problematic consequences (Arneson 1999: 95–97). I will not discuss that argument explicitly, since it is not among the arguments Rawls gives to justify FEO. However, I note that the argument is based on the assumption that the Priority View could work as a principle of anti-discrimination. I will question the latter assumption throughout Section 5. Note also that Arneson (1999: 104) finds the ideal of equal citizenship the only cogent argument for FEO.
The differences in citizens’ moral powers do not, as such, lead to corresponding differences in the allocation of primary goods, including the basic rights and liberties. Rather, the basic structure is arranged to include the requisite institutions of background justice so that citizens have available to them the general all-purpose means to train and educate their basic capabilities, and a fair opportunity to make good use of them. (Rawls 2001: 171)

For the purposes of justice as fairness, Rawls conceives of citizens as having two moral powers. They can act from fair principles of cooperation and they can act rationally; they have the capacity ‘to have, to revise, and to rationally pursue a conception of the good’ (Rawls 2001: 19). They are equal because they all have ‘to the essential minimum degree the moral powers necessary to engage in social cooperation over a complete life and to take part in society as equal citizens’ (Rawls 2001: 20). Citizens, moreover, are free in two respects. They view themselves and others as being able to revise their conception of the good, their goals and values, and ‘they regard themselves as being entitled to make claims on their institutions so as to advance their conceptions of the good (provided these conceptions fall within the range permitted by the public conception of justice)’ (Rawls 2001: 23). This moral conception of persons is the background against which Rawls claims, in the quote above, that differences in moral powers do not lead to a justification of different levels of primary goods, such as spending more resources on the education of the talented. Instead, and this is the meaning of the second sentence of the quote, opportunities, for these free and equal citizens, are to be afforded on an equal basis on the assumption of equality between citizens.

The opportunities needed to achieve equal standing as a citizen are to be distributed equally, which in turn probably means that more resources will be spent on the less talented than on the very talented. This also explains Rawls’s (2001: 174) insistence on general healthcare as a requirement of FEO. The ideal of equal citizenship explains why it is important that people have the opportunity to develop their talents, but it also puts limits on the requirement of equality in opportunities. Differences that do not have to do with equal standing fall outside the scope of FEO. David
Edmonds (2006) asks how we can make a distinction between skin color and toe size, when it may well be the case that the group consisting of people with small toes is worse off than other people. What is the difference between this situation and one in which the group of people with dark skin color is worse off? The Rawlsian answer is that toe size does not affect anyone’s social standing, whereas skin color has done this all too often. The core aim of FEO is to secure the equal standing of citizens, and its measure of unjust discrimination is this ideal.

A further, and even more fundamental reason to interpret FEO along these lines, can be found in Rawls’s account of how justice as fairness is an egalitarian theory. There are many ways in which equality can be seen as valuable, but Rawls opts for what he sees as Rousseau’s approach: ‘the fundamental status in political society is to be as equal citizenship, a status all have as free and equal persons’ (Rawls 2001: 132). He continues by spelling out the way in which equality is integral to justice as fairness and what this means for the interpretation of the theory:

Equality is present at the highest level in that citizens recognize and view one another as equals. Their being what they are—citizens—includes their being related as equals; and their being related as equals is part both of what they are and of what they are recognized as being by others. Their social bond is their public political commitment to preserve the conditions their equal relation requires. (Rawls 2001: 132)

What are these conditions? They are the social conditions of a society well-ordered by the two principles. This means that the principles must be read as a way of spelling out principles of justice from an ideal of relational equality. The core of justice as fairness is the ideal of relational equality, which is given by the account of free and equal citizenship. FEO must be interpreted as having the aim of giving expression to this ideal of relational equality. An illustrative way of explaining this abstract ideal is to look into how Rawls uses

---

2 Edmonds’ book is a good first introduction to the issue of discrimination. Two papers that also introduce important issues are Alexander 1992 and Wasserstrom 1977.

3 Several related versions of this view, but not Rawls’s theory, are criticized in Arneson 2013b.
the idea of slavery to explain the ideal of the citizen. Rawls says ‘[s]laves are, so to speak, socially dead: they are not recognized as persons at all’ (Rawls 2001: 24). This is not to say that we cannot care about the welfare of slaves, but they are not counted as sources of valid claims. Their moral powers do not count socially as a reason for equal standing and their conception of the good is not taken into account (expect perhaps indirectly by their holders). The citizen is the opposite of the slave. The ideal of the citizen is the ideal of independent persons.

Let us now turn to the relationship between FEO and the difference principle. Rawls aims to develop a theory of pure procedural justice and says of FEO that ‘[u]nless it is satisfied, distributive justice could not be left to take care of itself’ (Rawls 1999: 76). The goal is a situation where the distributive result of social cooperation is a function of individuals’ choices against a fair background. Without fair opportunity, citizens would have valid complaints of unfairness because they (as individuals) would not have equal prospects in deciding how to conduct their lives. If the difference principle were only constrained by a principle that says that positions should be open to all—if instead of democratic justice we opt for what Rawls dubs natural aristocracy—the charge from Rawls would be that this would carry with it the inequitable idea of noblesse oblige (Rawls 1999: 64). This is a serious issue for Rawls; his theory is devised to be suitable for a democratic society, and if the charge of bringing in aristocratic relationships and values, which conflicts with the ideal of free and equal citizens, into the core of justice as fairness succeeds, it would amount to a knock-down argument. This again illustrates the absolute centrality to Rawls’s project of the free and equal citizen: he is searching for the proper theory for citizens so conceived.

The relationship between the difference principle and FEO is, in short, that they are parts of a single principle based on the ideal of free and equal citizens. FEO is the part of this principle that regulates how the basic structure affects talents and ambitions, whereas the difference principle regulates the inequality that could result from differences in (natural) talents even after the opportunity principle has done its work (Rawls 1999: 87). Even if the more talented would have a better chance of attaining some coveted
position, the gains that they would have from this would be constrained by
the difference principle. If there are people who have better prospects of
becoming, say, judges, due to their innate talent, then FEO says that they
should have better chances to occupy such a position than those who
have less native talent for it. The difference principle then regulates
the structure of salaries and wages in society (as well as the tax system).
It is also important to see that it is not only the difference principle
that has egalitarian distributive implications. The first principle
demands the fair equal value of the political principles, whereas FEO
says that to the extent that inequality in terms of income and wealth
threatens equality in terms of opportunity, there ought to be redistribu-
tion. In this sense, FEO constrains the kind of inequalities that the differ-
ence principle would allow.

3. The ‘Too Weak’ Charge
Arneson’s first charge is that FEO cannot deal with the problem of stunted
ambition brought about by unfair traditions and cultural forces.

Any two persons of equal talent and ambition will have the same prospects of
success in competition for positions of advantage. However, this characteriz-
ation of the society is compatible with a further, disturbing description: all indi-
viduals are socialized to accept an ideology which teaches that it is
inappropriate, unladylike, for women to aspire to many types of positions of
advantage, which are de facto reserved for men, since only men come to
aspire to them. Any man and woman with the same native talent and ambition
will have the same prospects for success in the society we are imagining, but the
rub is that individuals’ ambitions are influenced unfairly by socialization.
(Arneson 1999: 78)

The first thing that strikes one when reading this is the question of how it is
possible to both have equal opportunities and not have equal opportunities.
The answer must be that FEO is too weak, because it only deals with the
design of institutions, and hence is silent on the issue of socialization.
According to this interpretation it would deal with influences on opportu-
nity other than those from social contingencies. Therefore, the argument
would go, you could have equal opportunities in the domain of this
narrow version of FEO, and yet have unequal prospects, all things con-
sidered. The charge of stunted ambition can only be an argument against
FEO if justice as fairness does not deal with socialization.

If the narrow interpretation needed to make the argument of stunted
ambition argument were correct, then it would make the fact that Rawls
says the following in the course of arguing for the second principle
appear rather strange: ‘There is no more reason to permit the distribution
of income and wealth to be settled by the distribution of natural assets than
by historical and social fortune’ (Rawls 1999: 64). The outcome of the
natural lottery is just as arbitrary from a moral perspective as are the
results of the social lottery. A basic criterion of adequacy on a theory of
justice is that it does not give natural or social contingencies free reign.
The idea behind Rawls’s second principle is to have FEO deal with the
social lottery and the difference principle with its natural counterpart. We
should prefer an interpretation of FEO that makes sense of this fundamen-
tal point for Rawls.

So cultural forces are within the domain of FEO, but one might ask
whether perhaps their effects on ambitions are not. Rawls takes the basic
structure as his subject

because the effects of the basic structure on citizens’ aims, aspirations, and char-
acter, as well as on their opportunities and their abilities to take advantage of
them, are pervasive and present from the beginning of life. (Rawls 2001: 10;
cf. 56–57)

This seems to imply that stunted ambition should be high on the list of pri-
orities for justice as fairness, since FEO deals with the cultural development
of aims and aspirations. It also suggests the usefulness of a term Arneson
coins in a later paper on discrimination. There, he says that what Rawls is
really concerned with is ‘prototalent’ (Arneson 2006: 795), which seems
to me to much better capture the core idea than plain ‘talent’. FEO deals
with how potential talent is turned into actualized talent. Again, we
should prefer an interpretation of FEO that makes sense of the claims
that Rawls makes.
We have seen that Rawls is concerned with class—‘same prospects of success regardless of […] social class of origin’ (Rawls 2001: 44)—but does this mean that he would be concerned with gender as well? There are textual indications that Rawls at least believes that he can address those issues. First, in his writings after Theory he has included sex explicitly as something that the veil of ignorance covers for the parties of the original positions (Rawls 2001: 15). Second, he has discussed the issue of justice between the sexes:

Moreover, to establish equality between men and women in sharing the work of society, in preserving its culture and in reproducing itself over time, special provisions are needed in family law (and no doubt elsewhere) so that the burden of bearing, raising, and educating children does not fall more heavily on women, thereby undermining fair equality of opportunity. (Rawls 2001: 11)

At this point, we have seen that there are three aspects of Rawls’s work that point in the direction of taking seriously socialization as it affects the formation of ambition with regards to gender: (1) the argument to the effect that the outcomes of the two lotteries cannot be consider morally justified, (2) a focus on the development of aims, and (3) direct textual evidence concerning gender equality. It seems very difficult to interpret Rawls as failing to see the kind of problem that Arneson claims that he misses, especially so when the theory takes as its starting point an ideal of relational equality. Moreover, the two first points in themselves seem to indicate that stunted ambition is not consistent with justice as fairness. FEO should be interpreted in such a way that it includes the development of this kind of development of ambition under its domain.4

In fact, it seems that Arneson does not find this a very telling argument either, because he ends the section devoted to this issue by saying that he has

4 Arneson’s argument is about the substantial content of FEO, which makes it distinct from the criticism of Rawls advanced by G.A. Cohen (1997). Cohen claims that the scope of application of the difference principle is too narrow and that it must be complemented with an egalitarian ethos. I believe that both arguments fail for similar reasons, namely that they do not take into account the full set of theoretical resources within Rawls’s theory, but since it falls outside the scope of this paper, I shall not try to show that this is so for the case of Cohen.
not done much more than put a label on this problem, after which he proceeds to suggest an amendment to FEO:

(a) that any two persons with the same talent and the same ambition should have the same prospects of competitive success and in addition (b) that the education and socialization processes that influence the formation of individual ambitions are unmarred by bigotry and unfairness. (Arneson 1999: 79)

If my arguments in this section are correct, this is not an amendment. FEO already contains the important aspects that Arneson claims it lacks in the original formulation. The Too Weak charge does not succeed.5

4. The Meritocracy Charge

The Meritocracy charge is in a way the most troublesome for Rawls. If it is correct, then he has incorporated, in the heart of his theory of justice, a principle that violates some of the core commitments of justice as fairness. A meritocratic FEO would disregard the argument from the two lotteries, undermine the emphasis on equal status and undercut the difference principle.

4.1 FEO and the Norm of Meritocracy

How does Arneson argue that FEO is meritocratic?

In a nutshell, the objection against Rawlsian Fair Equality is that it incorporates a compromise with the norm of meritocracy, the principle that holds that other things being equal, those who are naturally more talented and develop their talents to higher excellence levels should enjoy greater prospects of good fortune in life. Meritocracy says that positions that confer social advantages should be given to the most qualified, the standards of qualification being appropriate to the context. (Arneson 1999: 85)

The first sentence of the quote (which I shall call M1) says that the talented, if they decide to develop their talents, should be rewarded for being talented. This can be read as implying an unappealing desert idea of justice. The second sentence (M2) says that positions should be given to

5 Notice that the prioritarianism that Arneson (1999: 91–92) favors would take issue with the amended version of FEO that Arneson also favors.
the most qualified. It is important to keep these two senses of meritocracy apart. To give an example, even if one were committed to M2, one would be under no intellectual obligation to believe that the most qualified should be rewarded for the job they do in any particular way. There are then two different concepts of meritocracy at work in this quotation. There is nothing inconsistent in believing both that positions should be awarded to the most qualified and at the same time thinking that better prospects of fortune should not. It is not a logical mistake to think that all wages and salaries in society should be equal and that the most suitable candidate should get a given position. And, as one would suspect, Rawlsians are only committed to M2, which hardly qualifies as a real meritocratic principle (more on this below). If Rawls had been committed to actual meritocracy as well, it would have been an utter mystery how he could have argued for the difference principle, which under the right empirical circumstances demands strict equality, which would rather obviously be a complete violation of M1. Justice as fairness would incoherently imply that incomes should respect the difference principle and, simultaneously, that that income distribution should vary with merit. It does not.

So how does Arneson argue that Rawls must be committed to something like M1? He claims that Rawls is a meritocrat in two ways:

Here there is a choice between two states of affairs, (a) some equally talented and ambitious people are not made equal in their competitive prospects when this discrepancy could only be eliminated at the cost of making some worst off people even worse off, and (b) all equally talented and equally ambitious people are made equal in competitive prospects even though achieving this equality makes some worst off people worse off than anyone would otherwise need to be. Rawls’s priority rankings would have it that justice prefers b over a. Fairness to talent trumps fairness to the worst off in Rawls’s system. (Arneson 1999: 86)

We must keep the question of whether justice as fairness is meritocratic separate from the question of whether its priorities are wrongheaded. Even if we find that Rawls’s theory is not meritocratic, we might still find its priorities unfair. It is always the case that an anti-discrimination principle could be costly to the least advantaged. To avoid this result one would either have
to dispense with any anti-discrimination principle at all or to have it lexically subordinate. I will postpone discussion of this issue – the Too Strong charge – until the next section, and will instead continue, first, to make two points designed to show that the Meritocracy charge misses the mark (but leave open the question of whether FEO is too strong or not). As a second step, we will investigate the reasonableness of the priority ordering of FEO and the difference principle.

Since Rawlsians are not committed to M1, the aim of Arneson’s argument here must be to show that even M2 is meritocratic. To deny M2 would mean that one is committed to the claim that offices and positions should not be awarded on the basis of who is the most qualified. This suggests that Arneson’s argument works with a much too inclusive definition of meritocracy. Even such a historically salient opponent of meritocracy as utilitarianism satisfies it. As an example, consider rule utilitarianism. In a society governed by that doctrine the economy would be organized according to the system of rules that maximizes utility in the long term. Opportunities for employment, etc., would then be organized so as to further this goal maximally. What kind of employment policy would there be in such a society? Should the person who is the most qualified, in line with the optimal moral code, be hired, or the person who promises to deliver the maximum amount of utility in the situation at hand? Clearly, the rule utilitarian would be committed to the former. But this would then mean that a theory that is committed at its core to the denial of any grounds for distributive justice other than utility is, as it turns out, meritocratic. *Mutatis mutandis*, the same argument could be made for weighted utilitarianism, i.e., prioritarianism. I think that this shows that Arneson’s charge of Meritocracy fails—at least if interpreted as an argument based on the problematic aspects of meritocracy. M2 does not incorporate the troublesome norm of meritocracy. As Norman Daniels (1978) has shown, the idea of M2 (roughly) is compatible with a whole range of normative theories, most of which deny that distributive justice should be based on considerations of desert.

However, there could still be a case to be made for M2 being meritocratic, depending on how you spell out the rider ‘the standards of
qualifications being appropriate to the context’ from the quotation above. Doing so could show how utilitarianism, be it weighted or unweighted, is not meritocratic, whereas justice as fairness is infected with the meritocracy virus. This raises one of the most pressing problems in the moral theory of discrimination: how can we make a justified distinction between morally problematic and morally unproblematic forms of discrimination (cf. Altman 2011: ch. 5). FEO does this through the use of the ideal of free and equal citizenship. This ideal is distinct from the ideal of meritocracy.

At one point Arneson asks what could possibly be morally troublesome with having an African-American actor portraying Martin Luther King Jr. or criticizing situations in which quarterbacks in American Football are predominantly white (Arneson 1999: 79–80). The ideal of the free and equal citizen provides answers to these questions for justice as fairness. Discrimination is unjust when it is inconsistent with the achievement of the ideal of society as a cooperative venture between free and equal citizens. To choose an actor to portray a historical person on the basis of physical similarity does not seem to indicate a lower moral status of those actors who look dissimilar. Returning to professional sports, it was a good day for social justice when Jackie Robinson started playing major league baseball, even if this had no effect whatsoever on the economic prospects of the least advantaged of American society at the time. Likewise, we might find it problematic that quarterbacks are predominantly white and consider it an improvement if the pluralism of American society became reflected in this group of sports heroes. This indicates that justice as fairness gives plausible answers to the question of how to identify discrimination, without implying M1, the core motivation not being meritocratic concern with fairness to talent, but the ideal of equal social standing.

4.2 Lexical Priority

Arneson takes issue with the idea that we should spend resources on education for the aim of advancing FEO rather than aim to directly benefit the least advantaged. We turn now to that argument. However, here we would do well to recall the egalitarian aspects of the first principle of
justice, which underpins the fair values of the political liberties, and the fact that FEO constrains the inequalities that are allowed by the difference principle. These two distributive demands in themselves indicate a higher level of material equality than exists in the Western societies of today. Given, then, the place of FEO in Rawls’s theory, the situations we are discussing are cases in which discrimination would be an efficient means to improve the level of income and wealth in such societies, and given the empirical tendency of equality of opportunity to work in tandem with efficiency, we might guess that this would not occur very often. In this context it is important not to be led astray by visions of the kind of abject poverty that is much too prevalent around the world today.

We will start with Arneson’s most general claim, since the reply to this argument will have important implications for the other examples that he presents.

Fair Equality of Opportunity is defective in virtue of the lexical priority it is assigned over other justice values in Rawls’s scheme. Consider Fair Equality of Opportunity versus the Difference Principle. 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 1999: 81–82)

Now, we can, following Thomas Pogge’s (1989: 127) reading of Rawls, think of lexical ordering in justice as fairness in two ways: as design priority or as implementation priority. Design priority means that in setting up—designing—the ideal of justice, fair equality of opportunity has priority over the difference principle. Implementation priority means that when drafting policies to achieve justice, we should first create policies that secure the basic liberties. Only then are we allowed to implement policies to secure fair equality of opportunity, and not until this ideal is fulfilled are we permitted to start trying to achieve a distribution in accordance with the difference principle. We would not be allowed to spend any money on the poorest people in society until we had achieved equal opportunity. This is a bad
and rather irrational idea. It is also quite difficult to get this interpretation of lexical priority to cohere with the idea of the second principle being one principle. The basic puzzle that this idea of priority creates is why we are committed in the name of justice to achieve less justice than we could attain. Why not maximize expected value in terms of justice, or why not be a prioritarian when it comes to the implementation of justice as fairness? However, when we design a theory for the basic structure, the lexical ordering seems a good idea: it leads us to a conception of justice that includes a good deal of our pre-theoretical intuitions. The distinction we should make, then, is between what justice consists of (the ideal) and how to get there (implementation), and I submit that Rawlsians need only be committed to the lexical priority in the former sense. This means that justice as fairness can agree with common sense, in that if the least advantaged are very badly off and the FEO problem is minor, it is better to direct the resources at hand to bettering the position of the least advantaged.

4.3 The Upper Middle Class Example

If we understand priority as design priority, it is still true that justice as fairness would use resources to improve equality of opportunity that could have been spent on raising the economic level of the least advantaged. In order to decide if this is reasonable or not, we need some standards of reasonableness. Here Arneson argues ‘that the Rawlsian equality of opportunity principles fail to capture the most plausible considerations that underpin the judgment that the imagined caste society is unjust’ (Arneson 1999: 87). It seems, however, that these most plausible considerations are somewhat in contradiction, since Arneson proposes two different

---

6 A chronological interpretation cannot be right. The priority rule would rule out any application of the difference principle and destroy Rawls’s argument for a property owning democracy. Any change in the basic structure would upset FEO.
7 Moreover, even if Rawls himself were committed to the strictest possible form of implementation priority of the kind Arneson criticizes, there is no reason why a Rawlsian could not reject that position. No important aspect of justice as fairness turns on holding such a view of priority.
8 The quote talks about caste societies, but I assume that the same strategy is in play in all the examples I discuss.
principles to make his case. The first of these is the principle of Wrongful Discrimination (WD):

> Discrimination that causes significant harm should be forbidden if prohibition will prevent the harm without giving rise to comparable harms [...]. Discrimination that neither expresses hostile attitudes nor causes significant harm to anyone should be regarded as morally permissible, other things being equal. (Arneson 1999: 89)

This principle consists of two different clauses. The latter clause is pretty straightforward: in order to be permissible, discrimination must not express hostility or be the cause of significant harm. It is less clear how the formulation about harm in the first sentence should be understood. WD should be taken to mean either that the important aspect of harm has to do with the net (‘without giving rise to comparable harms’), or it concerns the level of harm inflicted on persons one by one (‘nor causes significant harm to anyone’). On the first interpretation, WD is straightforwardly consequentialist: whether or not discrimination is wrongful depends on the net amount of harm created by the alternatives at hand. However, Arneson prefers the second interpretation: only significant harm is wrongful.9 I will evaluate the examples on the basis of this interpretation in the following discussion: if an individual person is significantly harmed by discrimination, the discrimination counts as wrongful.

Arneson also evaluates FEO from what he calls the Priority View.10 According to this view discrimination is either an efficient means to achieving the prioritarian goal of improving the lot of the least advantaged, in which case it is morally justified, or it is not an efficient means to this end, and thus qualifies as blameworthy discrimination. Rawls’s difference principle is one example of how such a view could be specified, but Arneson thinks that leximin is too strict, preferring to give more room to consequentialist considerations. He does not, however, state exactly how we should understand the Priority View. On this view harms need not be significant in order to qualify as wrong.

---

9 See the semi-covert attempt at refutation of WD in Arneson 2006, where Arneson, while seemingly defending WD, accepts the implication that the abolitionists were engaged in wrongful discrimination while fighting slavery.

10 In the end this is Arneson’s (1999: 96) preferred view of anti-discrimination.
Arneson aims to show that there are better alternatives for a principle of anti-discrimination than FEO; since the aim of this article is to defend FEO from Arneson’s charges, my goal will be to show that these alternatives are in fact worse. Moreover, it is obvious that the two weapons that Arneson wields will give contradictory suggestions of justice in quite a few cases—for instance, the fact that the harm clause and the consequentialist Priority View disagree over the importance of net harm. I have my doubts about the usefulness of this argumentative structure. If we have two cases—one in which FEO is shown to be defective by the Priority View although WD agrees with FEO, and another in which FEO is rebutted by WD but the Priority View gives the same answer as FEO—what does this show? It cannot very well show that FEO is wrong in both cases.

Let us now turn to an example that seems to show that FEO cares about fairness toward the well-off more than it cares about the situation of the least advantaged. The example deals with the idea that resources should be spent on developing prototalent equally.

Imagine that even though much has been done to implement Fair Equality, it remains the case that individuals born with the highest talent potential in affluent upper middle class households grow up with lesser prospects of success than individuals born with identical talent potential in the very richest households, when individuals in both types of households develop the same ambitions. (Arneson 1999: 82)

FEO implies that this situation should be rectified, whereas common sense would presumably take this as an example of discrimination that could be deemed innocent. However, notice what a strange example this is. We are invited to think about a situation where the only relevant injustice is that highly talented children of the upper middle class have worse prospects than those born to the very richest.11 But if that is the only injustice, then

---

11 Arneson (1999: 82–83) sharpens this point further with an example of a society where the group of the least talented consists of a single untalented person. For this person, Arneson claims, FEO will be automatically satisfied, since there cannot be an inequality in a group of one. This example rests on the mistaken narrow interpretation of FEO that was discussed in Section 2, and misses the point that the aim of FEO is to bring about equality of standing
it follows that all other very talented children have better prospects than upper middle class children. If this is the only thing that stops equal opportunity from being achieved, everybody else must have equal opportunities. The talented children of the least advantaged in terms of income and wealth have better prospects than do upper middle class kids. It is hard to conceive of a plausible explanation of this highly unusual structural feature, but let me suggest two possibilities: racism or that it is an unforeseen consequence of policies aimed to achieve some other goal. An example of the first explanation could be the anti-Semitism that led to a situation where admission of Jewish students to a certain Ivy League university was restricted. Both my intuitions and Arneson’s say that this admittance policy was unjust (Arneson 1999: 93). This then boils down to an example that shows that prioritarianism by itself cannot serve as a principle of justice, as it would not be able to pick out the injustice of anti-Semitic admittance policies. This indicates the need for something like FEO or WD. Prioritarianism is deeply unsuitable as a principle of anti-discrimination, since it condones any kind of discrimination as long as it serves as an efficient means to improving the situation for the worst off. The choice between the Priority View and FEO should be obvious.

What would WD say? It would find that anti-Semitism is based on a hostile attitude, which then would show that FEO gives the right answer. But if we assume that the reason for this inequality of opportunity is not racism, but some mishap in policy, it seems that a situation in which one group is condemned for no apparent reason to worse prospects of achieving its goals would be just according to the hostility rider of WD. It seems to me that in a society in which the two principles with all their egalitarian implications are in effect, the intuitive response would not be to acquiesce to a state of affairs that singles out your child as having less than equal opportunities than other children. But perhaps intuitions differ here. What the harm clause would imply depends on whether we should view

among citizens. FEO will not be automatically satisfied in the situation that Arneson envisions, and therefore the example cannot serve as a counter-example to FEO.
this situation as giving rise to significant harm or not. If there is no significant harm in this example, then the harm clause would not even find the university admissions policy wrongful. If there is such harm, then we have a case of injustice. In the cases where ending this discrimination would inflict little harm on other parties the harm clause would say that Arneson’s example describes an instance of clear-cut Wrongful Discrimination. We can conclude that FEO would find this state of affairs unfair, WD would call it discrimination under some circumstances, and that prioritarianism gives terribly wrong answers in important cases, and can, therefore, not capture the intuitions of common sense. This does not show that FEO’s priorities are wrong-headed. Instead, it illustrates the need of a principle of anti-discrimination in the FEO vein. Arneson has not succeeded in showing that justice as fairness is meritocratic, nor that it is unfair; moreover, the alternatives he has put forth are less acceptable than FEO. The argument that FEO should be rejected on the ground of meritocracy fails.

5. The Too Strong Charge

It must be admitted that there is something *prima facie* implausible about the idea that opportunities are lexically more important than the resources distributed by the difference principle. This feeling is hard to shake, and may even remain after one has been reminded of the egalitarian implications of the first principle and of FEO itself. This would be especially so if it could be shown that FEO is concerned with issues ‘that should be deemed innocent rather than wrongful’ (Arneson 1999: 109). This hunch is the starting point of Arneson’s Too Strong charge. I will try to show that the implications of FEO are not counterintuitive.

5.1 The All-Gay Workplace Example

Let us now turn to an example that aims to show that FEO cares about the wrong things.

One sign that Fair Equality overreaches is that it not only forbids Christian heterosexuals, the vast majority of a society, from excluding gays from the vast
bulk of desirable employment opportunities, but also and equally forbids a few
gays who want to form a communal workplace from establishing a single business
in which all employees are gay. (Arneson 1999: 88)

To answer this charge, we should distinguish between three different situa-
tions, starting with the situation in too much of the world today. In a
social world in which homosexual persons are treated as less valuable
human beings, FEO proposes finding an efficient means of achieving
equal social standing. Intuitively, a plausible candidate for a policy would
seem to be to allow all-gay firms and come down hard on heterosexual dis-
crimination towards the gay community. This seems to be what Arneson, in
the quotation, takes to be the right kind of answer, and FEO does not dis-
agree, since stopping homophobic discrimination would be a major step
on the way to achieving the equality in standing FEO aims for.

However, there are possible scenarios in which Arneson is right about the
implications of FEO. This would be the case if both types of firms would
undermine the equality and freedom of citizens. In a situation where no-
one hires a person with a different background or identity to him- or
herself, FEO, and common sense, would find this troubling. This does not
seem a decisive blow against FEO. There is a third possibility: that both
types of firms would be allowed by FEO. In an egalitarian utopia in which
each person’s equal moral status is secure, and hence Rawls’s two principles
are implemented, there would be no case against all-straight or all-gay work-
places. Neither would there be any obvious use for them if homophobia were
overcome. The basic point here is that what is important is the effect on citi-
zens’ moral standing as equals that decides if and then under which circum-
stances FEO would condemn or allow any specific practice. This does not
show that FEO fails to distinguish between trivial and important cases.

How do WD and prioritarianism treat this case? As far as I can tell, WD
would not allow the all-gay workplace, even if it were the most efficient
way of bringing about a society without discrimination against sexual min-
orities, if this workplace were grounded on the basis of well-motivated hos-
tility toward a surrounding homophobic society, or if it significantly
harmed any heterosexual member of that society. If no such hostility or
harm were the case, WD would allow it. Arneson directs us to read the example as describing a situation in which there is no significant harm, in which case WD should find the situation unproblematic. However, if all gay workplaces were shown to be efficient means of achieving equality, WD would still not allow them if they harmed any of the heterosexual persons. WD is prone to make itself guilty of the same difficulty that Arneson accuses FEO of. The prioritarian would, as an example, ban the all-gay workplace, while allowing majority discrimination on the basis of sexual orientation, if it were the case that Christian charity depended on a tight-knit social structure based partly on homophobia. It is not at all clear that either WD or prioritarianism gives more intuitively compelling answers to question of discrimination than FEO. On Arneson’s own terms they seem to give exactly the wrong answers in some circumstances.\(^\text{12}\)

5.2 Skiing, Culture and Statistical Discrimination

Let us turn to an example that Arneson explicitly takes to show that even if FEO is based on the status of equal citizenship it is too strong in two different ways. This case concerns something seemingly trivial, but it raises some important issues.

A person of Asian ancestry who is very skilled at winter sports and just as desirous of gaining wintry jobs as a comparably skilled Norwegian will have lesser prospects of gaining these positions of advantage, given the statistical discrimination. A society ruled by maximin (more generally, by the Priority View) will restrict or prohibit statistical discrimination when it operates to lessen the prospects of the worse off, but not otherwise. But there is no reason to think that statistical discrimination per se will squash some citizens’ self-respect and render the status of equal citizenship for all a sham as it applies to them. A similar point holds if something in the culture of Chinese-Americans renders it the case that two children, one Chinese-American and one Nordic-American, who are born with equal winter sports potential talent, and who develop equally strong winter sports employment aspirations, end up with different prospects of success in competition for these positions of advantage (without any offsetting advantages favoring the Chinese group). (Arneson 1999: 105–106, my emphasis)

\(^\text{12}\) The reasoning of the last couple of paragraphs also applies to the very similar example of stunted ambition based on religious beliefs (Arneson 1999: 78–79).
I turn first to statistical discrimination, but here too I must begin by making some preparatory remarks. First, this example seems to be about trivial things; however, if we let it take place in, say, some small town in Oregon in which Nike does all the hiring, we would have different intuitions. If wintry jobs were important and no Asians could secure such employment, then any common-sense support for the prioritarian position might well dwindle. Second, if we are forced to make a trade-off when deciding whether to support FEO, WD and prioritarianism, because they all are found to have drawbacks, then surely even the fact (if it is a fact) that FEO cares about unimportant things carries little weight in comparison with the fact that WD and prioritarianism get central aspects of discrimination, such as homophobia and anti-Semitism, wrong.

Statistical discrimination occurs when people use proxies, e.g., skin color or sexual orientation, as indicators of characteristics of the individual in question. If left-handed people tend to do worse at work than right-handed people, this could be held as evidence that an individual lefty will do worse than an individual right-handed applicant. Obviously, we almost always face decisions under risk, or even uncertainty, and the rational thing to do is to adjust our probabilities in light of the evidence, regardless of whether this is based on characteristics of groups or of specific individuals. This might seem morally troublesome but Arneson argues that it is not.

Of necessity, people always respond to an individual on the basis of limited knowledge of her traits: one sees bits of evidence in behavior and appearance and so on and extrapolates to the future. (Arneson 1999: 92–93)

I agree, of course, with Arneson that choice based on statistical evidence is an unavoidable fact, but I fail to see how this justifies affording statistical discrimination the status of a morally justified principle. For that to make

13 Lippert-Rasmussen (2007) argues that there is nothing inherently wrong with statistical discrimination. The argument he presents seems right, but the conclusion follows because he defines away any problem of discrimination that has not to do with choices being made on the basis of statistical evidence. It might be better to label such decisions ‘statistical choice’ rather than ‘statistical discrimination’. My argument above rests on the assumption that there is actual discrimination at play in cases of statistical discrimination.
sense, ‘statistical’ would have to work as some kind of moral get-out-of-jail-free card. Theft or murder continue to be blameworthy, even if we stick a ‘statistical’ ahead of them; the addition of decision under risk does nothing to make discrimination more compelling. It still means that we are not treating equal citizens as equal. One can, of course, avoid committing murder, whereas it is unavoidable that one will at times use proxies, but this only shows that sometimes we should not hold people responsible for such discrimination. It does not mean that discrimination is good, or even neutral.

These considerations lead me to make two points. First, while it would be costly to eradicate all statistical discrimination, FEO would not aim for this utopian goal. Its goal is full and equal status for all. The final goal of justice as fairness is to make statistical discrimination based on ethnicity, gender, etc., meaningless. What it finds problematic is statistical discrimination that undermines equal standing. It is quite implausible to think that the discrimination of one person, concerning something that in any contemporary culture is of little importance, such as skiing, will be very high on our list of priorities for the near future. But it can be an injustice. We need FEO to evaluate statistical discrimination. Second, if we do take statistical discrimination to be justified (at a fundamental level), rather than regrettable, we find ourselves without a position to say that it is an improvement in terms of justice when people find a way of working around such discrimination. It would be better if Asian children were not disadvantaged by their ancestry when applying for jobs, and we should strive to overcome statistical discrimination, not celebrate it.

How would WD treat statistical discrimination? Assuming that the discrimination in question does not amount to hostility, the issue must be decided by the harm part of the principle: it ‘will condemn this per se innocent discrimination if its net costs that fall on those disadvantaged by it are significant’ (Arneson 1999: 92). We must return to the question of how we ought to understand the concept of harm. One way would be to take harm to mean (negative) welfare, i.e., preference frustration. But this does not seem right, because the answer to whether there is discrimination in the
skiing case would then depend entirely upon how upset the Scandinavian and Asian would get if we tried to intervene or refrain from taking action. We would find this a case of discrimination if the Asian-American person became very upset over being statistically discriminated against, as long as the Nordic-American were only moderately upset if we tried to right the imbalance. The opposite problem is almost certainly more common, namely, that oppressed groups internalize low self-esteem so that it seems natural that they should be discriminated against. When this is the case, there would never or very seldom be any wrongful discrimination according to WD. If women ‘know’ that they will not get that promotion and come to accept this due to resignation, they will not have been discriminated against. The welfare approach cannot be right.

We must look for a more objective standard. We should investigate whether we could base our calculations on the shares of resources that prioritarianism would afford the people involved. But, as we have seen time and again, prioritarianism is not a theory that gives reasonable results when applied to the problem of discrimination. This is because the Priority View shares with its parent, utilitarianism, a certain blindness toward these issues. Clearly, the more emphasis that is put on the least well off, the harder it will be to devise a credible example in which—to use one example that has haunted utilitarianism—prioritarianism would institute slavery, but there are no guarantees that it would be impossible. Arneson is right when he says: ‘[a] society ruled by maximin (more generally, by the Priority View) will restrict or prohibit statistical discrimination when it operates to lessen the prospects of the worse off’. But it is also true, as Arneson points out, that a society ruled by the prioritarian view would implement stunted ambition in women if this served the interests of the least advantaged (Arneson 1999: 91–92). In a word, prioritarianism is too weak. It seems we must look for some other objective standard, and here the ideal of free and equal citizens is one salient and attractive candidate. When trying to make sense of WD as an independent principle, we are led back to FEO. There would then be no (significant) harm in this case, only if the hiring policy did not affect the standing of the Chinese-
Americans. Either WD is wrong, or it is close to identical with FEO. Neither alternative gives us reason to reject FEO.¹⁴

Let us now turn to the second problem that Arneson presents in the winter sports quotation, that of cultural discrimination. How should we understand this example? By some fluke two children with the same innate abilities develop the same aspirations, but something in their culture stops the Chinese-American from achieving her goals. In order for this to be a problem distinct from that of stunted ambition, we need to find a different interpretation of what ‘something in the culture’ means. I shall assume that it refers to something in Chinese-American culture, as opposed to Nordic-American or general American culture. I shall also assume that this aspect of the culture is not something that the child in question has internalized, at least not fully, because that seems too difficult to combine with the fact that he or she has developed the same level of aspiration as the other child. Something in the culture must then mean the way in which other members of his or her cultural group behave toward him or her. And this behavior is in turn explained by some cultural norm that says, or has as its consequence, that Chinese-American children should not aim to achieve such goals. FEO finds this problematic for the same reasons that it finds stunted ambition a serious problem. The actions of others constrain the individual’s possibilities to develop his or her talents on an equal footing with others.

Enough has perhaps already been said about the drawbacks of the Priority View. But before moving on I want to make one further point, namely that it is sometimes too strong on Arneson’s terms. For the sake of exposition, assume that leximin is the best specification of the prioritarian doctrine. Now, in the culture problem discrimination does not entail any off-setting gains for the Chinese group, while there is perhaps a slight gain for the Nordic-Americans. If this is the only difference between the groups, then leximin would count it as an improvement (in terms of

¹⁴ Obviously, there could be other ways of developing an objective account of harm. I will leave the development of alternatives to the opponents of FEO.
welfare) that the Chinese-Americans are given the opportunity to compete on equal terms. They are the worst-off and their welfare would be improved by implementing a policy that did not frustrate their preference for employment in the winter sport industry. In some cases the Priority View, also, is too strong.\footnote{The same kind of argument could also be made against the less stringent version of the Priority View that Arneson prefers, but since the point is more easily illustrated by using the leximin variety I have chosen to use that version in the argument above. The reason that I say, above, that the Priority View is sometimes too strong is that in a case in which neither Nordic-Americans nor Chinese-Americans were among the worst-off groups, but both were, e.g., middle-class, a maximin interpretation of the Priority View might remain silent with regards to the kind of case Arneson presents.}

How would WD evaluate this situation? Leaving aside the issue of harm just discussed (the same considerations apply here as during the discussion on statistical discrimination), let us focus on hostility. It does not seem that hostility is a very probable factor. A much more likely motive is parental love and affection. Just as the parents through the centuries have taught daughters out of love to aspire to ladylike and passive roles in society, the children’s parents and others members of the culture may have taught and gently pushed their offspring into ‘appropriate’ roles. As the analogy surely shows, our theory of justice would be too weak if it could not identify such a state of affairs as a problem. Even if opportunities are constrained out of motives like kindness, love, or altruism, they are constrained.

To sum up, I have argued that there is little that is counterintuitive about the implications of FEO in the examples Arneson presents if priority is not understood as implementation priority. In all of the examples presented—the gay workplace, statistical discrimination, and something in the culture—FEO seems to give intuitively reasonable answers. I have tried to show the difficulties in making sense of WD and how it sometimes gives very counterintuitive answers to questions of discrimination, unless it is understood as a general formulation of FEO. Prioritarianism has been found to be profoundly wanting as a principle of anti-discrimination.
6. Conclusion
Arneson asks ‘[w]hy is it morally acceptable to single out the untalented and herd this group into the bottom rung on the social hierarchy?’ (Arneson 1999: 94). I hope to have shown that FEO does not aim for unequal citizens due to differences in talents, but rather for equal citizens with different talents. Furthermore, I have tried to be very clear about how the goal of FEO, and justice as fairness generally, is to overcome social hierarchy, not to create it. More specifically, I have argued that FEO is not too weak, as it does apply to the problem of stunted ambition; that it is not meritocratic in the unappealing sense; and that correctly understood it gives intuitive answers to the questions raised by the examples Arneson put forth, while the competition falls short in this regard. FEO is alive and well.

Acknowledgements: This work has been presented at the Royal Institute of Technology, Stockholm, Sweden and the Chaire Hoover d’éthique économique et sociale at UCL, Louvain-La-Neuve, Belgium; thanks to those who participated. Special thanks to Anca Gheaus, Sara Belfrage, Dan Munter, Kalle Grill, Niklas Möller, Axel Gosseries and Mikael Dubois.

Funding: This research was funded by the Swedish Research Council [grant number 2013-2337].

References
In Defense of Rawlsian Fair Equality of Opportunity


