# Two Models of Equality and Responsibility

## Citation
Blake, Michael, and Mathias Risse. 2008. Two models of equality and responsibility. Canadian Journal of Philosophy 38(2): 165-199.

## Published Version
http://dx.doi.org/10.1353/cjp.0.0018

## Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:3710803

## Terms of Use

# Share Your Story

Accessibility

# Two Models of Equality and Responsibility

Abstract: Much recent political philosophy has focused on the role of responsibility within liberal-egalitarian theories of justice, and much of that has been very critical of Rawls's *Theory of Justice*. This debate overlooks a central distinction within the taxonomy of theories of justice. There exist at least two visions of how a liberal-egalitarian theory of justice can integrate considerations of distributive equality and, derivatively, responsibility. What differentiates these two versions is whether distributive equality is taken as immediately plausible or a direct expression of respect for persons (or a related notion), or whether distributive equality is derived only in the presence of additional claims about normatively relevant relationships among persons. These models imply distinct accounts of how responsibility may be integrated into accounts of liberal-egalitarian justice. This fact allows us to show that, contrary to what has motivated much of recent political philosophy, as far as responsibility is concerned, Rawls's approach is not already problematic on its own terms.

## 1. Introduction

1.1 Much recent political philosophy has focused on the role of responsibility within liberal-egalitarian theories of justice.[1] John Rawls's theory, in particular, has come in for criticism in virtue of its account of responsibility, which Rawls takes to be tied to his account of primary goods. Primary goods themselves have been rejected as the appropriate "currency" (or metric) of distributive justice, while Rawls's treatment of responsibility has been criticized as implausible or even inconsistent. These criticisms have given rise to much of the constructive work political philosophy has done after Rawls's *Theory* -- one may think of the work of Richard Arneson, G. A. Cohen, Ronald

---

Dworkin, and John Roemer, for whom a major concern is to make more room for a suitable notion of responsibility.[2] Their efforts are shaped by a distinction between "choice" and "circumstance:" individuals should possess distributive shares in accordance with their choices (for which they are responsible) and be compensated for disadvantages they have because of their circumstances (for which they are not.) These writers are sometimes called *luck-egalitarians* (cf. Anderson (1999) and Scheffler (2003)), but their concerns are more aptly emphasized by calling them, following Arneson, *responsibility-catering egalitarians*.

There are two accounts of how responsibility-catering egalitarianism relates to Rawls's *Theory*. The first emphasizes continuity: Rawls offers insights about the role of responsibility in social justice but leaves them underdeveloped, and it is up to others to fill the gap. The second account emphasizes discontinuity: to the detriment of his theory, Rawls fails to discuss responsibility adequately, and thus social justice must be reconceived to accommodate an appropriate notion of responsibility. The continuity reading appears in Kymlicka (2002), and a recent interview suggests that Dworkin also endorses it (cf. Pauer-Studer (2002)).[3] The discontinuity reading appears in Roemer (1996). We believe there are strands in Rawls (1999a) to support both readings, but on balance, as Scheffler (2003) also argues, the discontinuity approach is the better reading: the differences between a Rawlsian account of social justice and responsibility-catering

---

[2] Cf. Cohen (1989), Dworkin (1981), Dworkin (2000), Arneson (1989), (1990) and (1997), and Roemer (1996).

[3] This is a bit surprising, since Dworkin (1981) introduces the criticism of Rawls that his difference principle is insensitive to questions of responsibility in terms that fit the discontinuity reading. Brian Barry seems to be at least sympathetic to the continuity view as well: cf. chapter 9 of Barry (1995), especially pp 54-55.

egalitarianism are too big to support an account highlighting continuity. In this spirit, we must proceed to offering actual responses to the objections to Rawls's treatment of responsibility that motivate responsibility-catering egalitarianism.[4]

1.2 Much of this debate overlooks a central distinction within the taxonomy of theories of justice. There exist two different visions of how a liberal-egalitarian theory of justice can integrate considerations of distributive equality. On the first account, distributive equality - of whatever it is that is distributed – is a necessary implication of the foundational moral commitments of a theory of justice. We will refer to this type of theory as a *direct* theory of distributive justice. There are at least two possible ways in which such a theory might be constructed. The moral relevance of distributive equality might be understood in axiomatic terms: for such a theory, the distributive equality is the foundational commitment of the theory, and there are no more basic claims from which distributive equality is actually derived. Such accounts are committed to distributive equality at the

---

[4] There are related contributions in the literature. Lippert-Rasmussen (2001) and Scheffler (2003) discuss similar themes but their discussions of criticisms of Rawls focus on Dworkin. Scheffler (2005) argues that luck-egalitarianism is misguided because their reasons for wanting to integrate responsibility into egalitarian theory conflict with the concerns that motivate egalitarianism to begin with. In our terminology, Scheffler could be straightforwardly classified as an indirect theorist who argues against direct theorists that they have lost sight of the importance of morally important relationships before the background of which concerns of distributive equality become relevant. (A similar point holds for Anderson (1999).) Schaller (1997) discusses Arneson, but his concerns are somewhat different. Ripstein (1999) can be understood as systematically developing the implications of the Rawlsian "division of responsibility" within legal contexts (while his chapter 9 and Ripstein (1994) also look at the wider distributive context). Our contribution is twofold: first, we develop a distinction between different ways in which the notion of responsibility can be integrated into a theory of distributive equality. And second, we apply this distinction to objections against theories such as Rawls's based upon a supposedly inadequate notion of responsibility. We focus on Arneson and Roemer because their arguments have not yet received adequate commentary.

axiomatic level. The currency, or metric, of distributive justice is such that distributive equality is by itself a demand of liberal egalitarianism. More often, however, the moral relevance of distributive equality is understood as an immediate implication of some more foundational notion of equal respect for persons, or some similar notion of moral equality. Such accounts are committed to distributive equality as a direct derivation of their moral axioms. While equal distribution of the currency of distributive justice is not here an axiomatic demand of liberal egalitarianism, equal distribution follows, without additional assumptions, from the conception of equality developed. Treating people as moral equals, for such theories, implies the provision of equal distributive shares. For our purposes, what links such theories and axiomatic theories is the fact that in both cases distributive conclusions may be derived from abstract moral egalitarianism. We will refer to all such theories as *direct* theories of distributive justice.[5]

The alternative account might be referred to as an *indirect* theory of distributive justice. Here, distributive equality may be derived from an account of equal respect of persons or some other account of equal personhood, but *only* in the presence of (and essentially through, in a sense explained later) additional claims about normatively

---

[5] Our use of "direct" and "indirect" should not be confused with the use of these terms in the analysis of consequentialism. In that context, this terminology is used to distinguish conceptions of consequentialism that assess the objects of evaluation (normally actions) in terms of their consequences from conceptions that assess them in terms of rules or motives that in turn have good consequences. The former set of conceptions of consequentialism is sometimes called "direct" (and act utilitarianism is a common example), whereas the latter set is sometimes called "indirect" (rule utilitarianism being a common example). (Cf. article on "Consequentialism" by Walter Sinnott-Armstrong in the on-line Stanford Encyclopedia of Philosophy, http://plato.stanford.edu/entries/consequentialism/). We do not claim any link between the notion of directness in the two contexts; rather, the notion of a "direct" theory of justice refers only to the process of logical deduction, not – as in the case of consequentialism – the level of specificity at which moral conclusions are derived.

relevant relationships among persons. The particular case we will be interested in is one in which the additional relationships are ones of shared citizenship.[6] Such accounts of liberal egalitarianism are committed to *indirectly derived distributive equality*. On this third model we cannot automatically infer any (even prima facie) injustice from distributive inequality, as we can on the other two. Distributive inequality may merely reveal that the relevant relationships do not hold. What concerns us in this study is the difference between direct and indirect models of distributive justice. The former defend distributive equality without making this demand dependent on the presence of normatively relevant relationships that not all persons may (and in practice will not) share, whereas the latter do appeal to the presence of such relationship. We will refer to both models characterized by axiomatic distributive equality and accounts characterized by directly derived distributive equality as *direct* accounts of equality, and refer to accounts of the remaining type as *indirect* accounts.

Which model we adopt matters for assessments of distributive inequality. In particular, these models imply distinct accounts of how responsibility may be integrated into accounts of liberal-egalitarian justice. This fact allows us to defend Rawls against his critics, by demonstrating that the criticisms described above use a theory of responsibility appropriate only to an account of distributive equality alien to Rawls's theory. To wit: Rawls's, we will suggest, is a theory that endorses indirectly derived distributive equality,

---

[6] We do not deny, however, that other such relationships might exist; a fully specified comprehensive liberalism might, for instance, specify certain sorts of distributive consequences applicable only within the family. Our present focus, however, is on the particular case of political societies, and we will therefore restrict our attention to those theories such as Rawls's political liberalism, which focus on distinctively political relationships such as shared citizenship.

whereas the critics to be discussed hold a theory in which distributive equality is derived directly. We confine ourselves to liberal-egalitarian theories of justice because the debate we discuss is internal to that family of theories. We confine ourselves, further, to explicitly *egalitarian* theories of distributive justice, although what we say applies also to variants such as prioritarianism. The distinction between those three approaches to equality carries over to such theories. One cannot avoid the conclusions we draw from this distinction simply by claiming to be a prioritarian *rather than* an egalitarian.

We do not here defend an approach to distributive equality. We do think the approach that derives equality indirectly should be chosen, but our concern is to develop the distinction and to show that it allows us to rebut certain criticisms to Rawls's account. This moves the dispute to an assessment of which approach to distributive equality should be preferred, which is where the debate should be. But as far as responsibility is concerned, Rawls's approach is not already problematic on its own terms.

1.3 Section 2 describes the models of equality we seek to articulate. We show how the differences between these models determine the role of responsibility within the chosen theory of justice. Section 3 discusses the Rawlsian notion of a "social division of responsibility," a conception in virtue of which Rawls's is a model in which distributive equality is derived indirectly. Sections 4 and 5 apply this distinction to some of the disputes described above. We seek to show that some of Rawls's critics employ a view of responsibility that is untenable in light of Rawls's commitments and cannot serve as a basis for objections. The critics, that is, are committed to a notion of responsibility available only in direct models of distributive equality. Section 4 establishes these

6

conclusions with regard to Arneson's criticisms. Section 5 does so with regard to Roemer's. Roemer is not only concerned to show that Rawls's view of responsibility is inconsistent, but also that egalitarians should drop the veil-of-ignorance device. We hope to show that Roemer is wrong.[7] Section 6 concludes.

## 2. Direct and Indirect Models of Distributive Equality

2.1 Logically speaking, there are three models of the place of distributive equality within a liberal-egalitarian theory of justice. While we discuss them, keep two points in mind. To begin with, in each of these models the goods at issue are whichever goods a theory of justice distributes. For none of these models does this mean necessarily that those goods offer a theory of the "the good" per se – though it might, depending on the account. Moreover, we do not intend for some of these models to capture a moral notion of equality whereas others capture a political notion. The concern is with a classification of theories of justice; how the words "moral" and "political" are to be used depends on the details of these theories.[8]

It is possible, first, simply to stipulate the relevance from the standpoint of justice of an equal distribution of some good – be it welfare, opportunity for welfare, capability, or some alternative conception. If our conception of equality within a theory of justice is spelled out in terms of the equality of some good, the distribution of that good becomes

---

[7] Roemer (1996) contains the most extensive critical discussion of Rawls, whereas Roemer (2002) (drawing in Roemer (1996) and Roemer (2001)) argues that egalitarians should not use the veil-of-ignorance device.

[8] That said, we will talk about moral and political notions of responsibility within the context of the Rawlsian theory and its critics, but by that time we will have introduced enough material about the Rawlsian theory for this to be appropriate and clear.

relevant for an assessment in terms of justice without further argument. On this account, the fact that two individuals differ in their holdings of this good bears immediately on our evaluation of their situations. Unless defeated by opposing considerations that the respective theory of justice may also acknowledge, the conclusion that an injustice exists is immediately derivable from this situation. An unequal distribution of the relevant good is *prima facie* an injustice. We say that such accounts are committed to *axiomatic distributive equality.*

In the second model, it is an ideal of equality of persons, most plausibly equal respect and concern for persons, that calls for the equalization of the distributive goods identified by the theory as relevant for the purposes of justice. Distributive equality is not stipulated, but derived from a normative account of personhood. This model is committed to *directly derived distributive equality*. While the first and second accounts are logically distinct, we do not further distinguish between them. An account of distributive equality committed to axiomatic distributive equality without endorsing any notion of equality of persons from which the distributive equality can be taken to be derived, though logically possible, would be rather peculiar. We discuss both types of accounts together under the heading of *direct models.*

The third model sees distributive equality as relevant only in a mediated way. The link between equal concern and respect, or any other notion of equality of persons used in the respective account, and the distributive goods identified is not through an inference that requires no additional assumptions; rather, equalization of these goods is demanded only in the presence of and essentially through additional normatively relevant relationships among the respective individuals. While distributions of the chosen metric

8

of goods are not irrelevant from the standpoint of justice, their importance must be established through additional arguments showing that the relevant relationships hold. We cannot, from the sheer fact of an inequality in holdings of the good at stake, infer even a *prima facie* injustice. This model is committed to *indirectly derived distributive equality*. For brevity's sake we talk about the *indirect* model of distributive equality.[9]

The qualification "only in the presence of and essentially through additional normatively relevant relationships among the respective individuals" requires elaboration. What this formulation is intended to make sure is that direct theorists could not simply become indirect theorists by adopting the condition that either the axiomatic postulation or direct derivation of distributive equality only applies to a certain range of people (say, those living within a certain jurisdiction). If this were possibly, there would not be much depth to our distinction. That qualification makes sure that the presence of those normatively important relationships is actually *needed* for the derivation of distributive equality. The next subsection discusses some theorists and thereby should illuminate the distinction between direct and indirect models, and especially make clear the relevance of this qualification.

2.2 To illustrate, consider first Rawls's theory of justice. According to Rawls, primary goods are not always the proper metric for the evaluation of equal concern and respect. They become the appropriate metric only through arguments demonstrating that they properly express egalitarian concerns within a particular context – the basic structure of a

---

[9] When we talk about "holdings of the good at stake," the good might, of course, be welfare. So our talk about "holdings" is by no means confined to material objects.

constitutional democracy.[10] The underlying notion of equality of persons by itself is not taken to have distributive implications. Rawls's arguments are designed to show that the application of such a general framework to a specific social context motivates the relevance of a specific metric of egalitarian concern – a metric which is given its relevance *by* that context and is applicable only *within* that context. This additional argumentative step marks Rawls's theory as an indirect model of distributive equality.

Direct theories have no space for this argumentative step. An example is a welfare-egalitarian variation of Peter Singer's brand of utilitarianism. Singer (1993) starts with a principle of equal consideration of interests, which he takes to capture the form of moral judgments.[11] Equal consideration of interests is taken to have immediate distributive implications, captured in terms of well-being, but for Singer these implications are utilitarian (maximizing well-being, rather than equalizing it). Neither the significance of well-being per se nor that of its distribution depends on specific social or

---

[10] The basic structure of society is the way in which the main political and social institutions of society fit together into one system of social cooperation. The two principles Rawls argues will be chosen in the original position (behind the "veil of ignorance") to regulate the basic structure are the following (in lexicographic order; cf. e.g., Rawls (2001), p 42): (a) Each person has the same indefeasible claim to a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all. (b) 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 (the "difference principle").

[11] We use Singer as an example because his moral theory is a particularly clear and simple vision of direct theorizing; we read him as a maximizing utilitarian in the tradition of Jeremy Bentham. As such, he represents a thinker for whom distributive consequences are directly the result of his axiomatic structure – namely, the welfare of all sentient creatures. Indeed, he has been particularly forthright in refusing to allow contextual relationships to enter into his calculations of justice; see Singer (2005).

political contexts, and much of Singer's work as a whole argues against assigning much relevance to such contexts. Were Singer a welfare-egalitarian rather than a utilitarian, his would be a paradigmatic direct account of equality within a theory of distributive justice. G.A. Cohen, Richard Arneson, John Roemer, and Ronald Dworkin are, on this analysis, direct theorists as well.

While we will not be concerned with Cohen and discuss Arneson and Roemer in sections 4 and 5, we should explain why we claim Dworkin as a direct theorist. Dworkin (2000) explains that his book "argues that equal concern requires that government aim at a form of material equality that I have called equality of resources" (p 3). Dworkin's theory of equality of resources attempts to explain how that sort of distributional equality is derived from equal concern and respect for people. This much seems to qualify Dworkin as a direct theorist *par excellence*, but he explains as well that "the principle of equal importance does require people to act with equal concern toward some groups of people in certain circumstances," and what characterizes these groups is that they are a "political community that exercises dominion over its own citizens, and demands from them allegiance and obedience to its laws… Equal concern (…) is the special and indispensable virtue of sovereigns" (p 6). In light of this comment he may seem to qualify as an indirect theorist instead, since his derivation of distributive equality applies only to people subject to the same sovereign jurisdiction.[12]

Yet now it becomes important that we define an indirect theorist as one how derives distributional equality not only in the presence of, but also *essentially through*

---

[12] This interpretation of Dworkin seems to inform Thomas Nagel's interpretation of his work.  See Thomas Nagel (2005).

additional normatively relevant relationships among the respective individuals. Dworkin's derivation of distributive equality proceeds independently of these sovereignty-considerations: restricting the scope of people affected by this derivation in this manner requires an additional and independent argument, which Dworkin offers. The fact that the relevant individuals are tied by being subject to the same sovereign jurisdiction (a constraint explicitly or implicitly built into many theories of justice that do not take themselves to have international implications) *does no work* in the derivation of distributive equality. This sets Dworkin apart from Rawls, in whose account it is crucial for the derivation of distributive equality that the affected individuals are free and equal citizens in a system of fair cooperation. For Rawls, it is indeed essentially through normatively relevant relationships that distributive equality is derived; for Dworkin it is not. Therefore, as far as the derivation of equality is concerned, Rawls is indeed an indirect theorist while Dworkin is a direct theorist.

The adoption of such a model has significant implications for the sort of reasoning that is appropriate for a range of questions, and hence also as to the substantive conclusions the theorist will reach. It is because of these implications that our classification is relevant. We will here consider three topics: reasoning about the *scope* of such theories will be distinct; about the *range of goods* over which it has evaluative jurisdiction; and about the notion of *responsibility* it takes as salient. The subject of responsibility will then occupy us for the remainder of this study.

2.3 To begin with, within an indirect account, it is straightforwardly plausible to restrict the *scope* of distributive justice to some set of individuals selected by the theory. The

contours of this set are determined by the normatively relevant relationships endorsed by the theory. Within the Rawlsian project, the relevant relationship is that of being jointly engaged in social cooperation. If the original position is, as Rawls insists, not a device to be employed every time there is the division of a good, but a heuristic device designed to show how the principles of justice animating a constitutional democracy might be derived, the parties to the original position will be limited to those subject to the coercive control of the political power of the state.[13] Generally, what is required of respectful treatment depends on what structures (cooperative or coercive) the individuals share.

This approach has implications in a variety of areas; perhaps the most important is that of *international justice*. For the kind of welfare-egalitarians for whom the fact of an inequality in welfare is directly relevant, the different levels of well-being in Western and developing economies are a *prima facie* injustice.[14] Unless some further factor can be introduced to defeat the conclusion that differing levels of welfare are illegitimate, such inequality gives us reason to condemn the world's current economic and social situation. Yet for Rawls the fact of an inequality in social primary goods becomes problematic as a matter of justice only within the context of a domestic political society, and only because equality of persons demands this equalization within this specific context.[15] Some Rawlsians have chosen to extend the scope of the original position past what Rawls

---

[13] Cf. Rawls (1993), pp 136-137.

[14] For a well-known example of such reasoning, cf. Singer (1972).

[15] For a justification of this reading of Rawls, see Blake (2001). Cf. Risse (2005) for an account of duties in the international arena that is Rawlsian in outlook.

himself regards as its appropriate space.[16]   Others – notably Rawls himself – have

rejected this argument; whatever duties of justice exist in the international community,

they will be distinct in form and content from the duties of domestic justice.[17]


2.4 Reasoning about the *range* of goods over which distributive justice extends, secondly,

will be distinct in the two models.  For the direct model, all aspects of life that influence

the distribution of the good in question are within the purview of justice: since equality of

a certain good is immediately relevant from the standpoint of justice, or is taken as an

immediate expression of equality of persons, whatever affects its distribution might

similarly be regarded as at least prima facie relevant from the standpoint of justice. It will

take independently plausible considerations to exclude some factors that affect that

distribution, considerations that would have to be consistent with whatever reasons the

direct theorist has for choosing the good whose distribution matters immediately and for

insisting on the equality of its distribution.

  We may take Anderson's (1999) hypothetical idea of compensation for the

sexually unattractive as instructive (cf. p 305). While some theories of justice reject such

an idea, direct accounts can only reject it for second-order reasons; they will invoke

considerations of stability, argue that such compensation risks endangering the just

distribution of goods in other spheres of human experience (perhaps by producing

---

[16] Cf. Beitz (1979) and Pogge (1989). We do not mean to suggest that the adoption of an indirect model by itself *commits* one to a view of international justice of the sort captured by Rawls (1999c). Our point is only that different reasoning will occur in direct and indirect models when it comes to international justice.

[17] Cf. Rawls (1999c).

additional evils such as resentment and shame among those so compensated), or that equality is a political virtue of sorts, one which is only saliently held over a specific range of goods (excluding sexual attractiveness).[18] Yet indirect theorists are able, without further ado, to restrict the scope of their theories. Indirect theories acknowledge the existence of normatively important relationships vis-à-vis which the goods whose distribution is to be regulated derive their importance, and it is in light of the nature of that relationship that the range of these goods is restricted. Within the Rawlsian project it is natural to restrict the range to goods created and distributed through cooperative activity. Since it is the creation of principles of justice to adjudicate this realm that is sought, goods outside this realm are not implicated within a direct account of distributive justice.[19]

---

[18] It is worth emphasizing that these are two different strategies. The first offers second-order reasons to defeat what is a prima facie valid implication of a direct egalitarianism; the second argues that the direct egalitarianism has a space of goods that does not extend to this particular form of value. The second strategy is one which is available to both indirect and direct theories; any particular putative good – such as sexual attractiveness – might fall outside the range of goods recognized in a given theory, whether indirect or direct. Our claim at present is that a direct theory can exclude such goods even when they are indeed potentially relevant from the standpoint of justice, by demonstrating that the individuals in question do not stand in the correct form of relationship to one another to make their claims valid. We are grateful to an anonymous editor for this Journal, for pressing us to be more clear on this point.

[19] It is possible, of course, that considerations of sexual attractiveness might fall within the purview of justice – if such "goods" can be represented as linked importantly to those goods relevant to political justice. We might imagine, for example, that the attractive have greater access to political power; in this case, even an indirect theorist would have to conclude that a maldistribution has occurred. Here, however, the maldistribution comes in the linkage between the good of political power and sexual attractiveness, rather than in the distribution of sexual attractiveness itself. The latter, for an indirect theory such as Rawls's, does not fall within the purview of distributive justice, and so the distribution of this "good" may be allowed to retain whatever shape it happens to take. We are grateful to an anonymous reviewer for pushing us to be more explicit on this point.

While the example of the sexually unattractive is fanciful, the difference in range extends to more serious cases. Kymlicka (2002) charges that Rawls is unable to justify his refusal to consider compensation for inequalities in *natural* primary goods (pp 70-72). He argues that Rawls's reasons for demanding the equalization (subject to the difference principle) of social primary goods translate to the sphere of natural primary goods; if we care about human well-being, surely we would care as much about the latter as the former. Yet the distinction between direct and indirect models demonstrates how this charge might be avoided. For theorists like Rawls, it is not well-being or life-chances that are directly relevant for justice. What *is* relevant is the project of justifying coercive authority to those subject to a state's jurisdiction. In this project, distributive shares may become relevant, inasmuch as the economy is part of the basic structure. In order to justify what we do to one another through the basic structure, we have reason to ensure equalization of the benefits and burdens the basic structure creates. We have reason to care about distributive shares of social primary goods, because such goods are ultimately the product of a political community in need of justification.

Natural primary goods, in contrast, exist prior to and (largely) independently of the basic structure.[20] They become relevant only when (and to the extent that) they are significant for the positions and offices the basic structure creates. On this analysis, there is nothing inconsistent in insisting that natural primary goods such as talents and

---

[20] We say "largely" because the influence that natural primary goods will have on individual life-chances will depend to a great extent on the arrangements of the basic structure. Moreover, the basic structure can be organized in such a way that it allows compensation for shortfalls in natural primary goods (e.g. through transfer payments to the disabled, or, alternatively, through the legal provision of mandatory access for the disabled to public buildings, etc.). On this, see, for example, Wolff (2002).

intelligence must not be allowed to influence distributive shares, barring justification

through the difference principle – while insisting also that inequalities in natural primary

goods are not to be compensated for by political authority. (What is more: we hope that

this discussion also makes it plausible that there is a normatively appealing vision behind

these views.) We ensure that intelligence and talents do not illegitimately affect social

primary goods without taking it as required that intelligence and talents themselves be

equalized or compensated for.[21] What Kymlicka understands as an inconsistency is the

result of a consistent reading of an indirect model of distributive equality.[22]

      This discussion highlights how using an indirect model of distributive equality –

mediated through the kind of reasoning that accompanies it -- has consequences for

conclusions about justice.   The differences in natural primary goods that Kymlicka

highlights are not trivial; differences in natural endowment can represent important

differences in well-being, even when such differences do not become relevant for

---

[21] It should be noted that Rawls briefly adverts, in *A Theory of Justice*, to a principle of redress, on which all natural inequalities might be taken as prima facie unjust and therefore deserving of compensation.  This principle, however, is not one Rawls consistently endorses; he identifies it only as one possible strand in our shared public understanding of justice, and argues that such a principle is best achieved through the more constrained conception of justice embodied in the difference principle.  As such, while the principle of redress might be understood as a nascent form of direct theorizing, it is difficult to ascribe such a principle a significant role in Rawls's thought.

[22]  It will be useful, at this stage, to recall the distinction between the continuity and discontinuity reading of Rawls that we briefly addressed in the introduction. One may say that it is precisely those passages in Rawls that support the continuity reading (and thus put Rawls in proximity to responsibility-catering egalitarians) that reappear at this stage and support Kymlicka's inconsistency charge. However, this discussion above should make clear how the discontinuity reading that we favor can accommodate those passages (in particular Rawls's talk about "moral arbitrariness" that seems to apply to social and natural primary goods indiscriminately).  That is, concerns about natural inequality do enter an account of distributive equality, but only in the manner sketched above.

positions and offices within the basic structure. The adoption of an indirect model represents a considered conviction that not all losses, and not all inequalities, matter for a theory of justice. We seek to equalize benefits and burdens produced by political community (with benefits and burdens not so produced entering through their effect on those that are), so that the authority of political government might be justified; we do not seek to equalize burdens and benefits themselves. Some losses are left to lie where they fall, and for others there will be a heavy burden of proof on those who argue that they should be in the purview of justice. My lack of talent will only be a matter of concern for a theory of justice once that lack begins to affect the goods I am able to acquire in the market; considered in itself, it is simply bad luck.

2.5 The third and for our purposes most significant implication of our distinction concerns different possibilities for integrating a suitable notion of responsibility into an account of distributive justice. Note first, however, that there is nothing internal to either direct or indirect accounts of justice per se that compels them to accommodate a notion of responsibility, so the reasons for wanting to do so must arise independently. We will see below how such reasons arise for Rawls's indirect account, so here we merely point out how they may arise straightforwardly on a direct account.

For illustrative purposes, suppose first our direct theorist is a welfarist. Welfarists make an individual's claims to distributive shares dependent on factors that do influence her welfare, but that many think should not bear on her distributive share. One key example is the "expensive tastes" problem. This problem involves an individual deliberately cultivating expensive tastes and thereafter asserting a claim for greater

distributive shares.  Here, however, it is intuitively inappropriate to cater to such desires; thus many theorists attempt to construct an account of responsibility in which the agent's deliberate choice to develop these tastes is taken as sufficient reason to deny them as sources of egalitarian claims.

Suppose now our direct theorist is a resourcist. By way of contrast with welfarists, resourcists face a fetishism objection: it is not the possession of resources as such that we should care about (at least not exclusively), but what resources do for people. Without any fault of their own, that is, in a manner that cannot be traced to any voluntary choices they made, some people have more needs than most (e.g., handicapped people), while others may have acted in ways that make them deserving of more than an "equal" share of resources.  Again, the notion of responsibility enters as a means of correcting the defects of an egalitarian distribution; here, the fact that the disability is not something for which the agent is morally responsible might be used to justify her greater claim upon resources.

Both of these direct theories therefore seem to require *corrective criteria* to be plausible as accounts of distributive justice. Considerations of voluntary choice provide such criteria. We can see, then, how it is tied to this notion of voluntary choice that the notion of responsibility may enter such accounts.  The notion of responsibility is employed to adjust and improve the egalitarian distribution of the theory in question. This brings consideration of responsibility into the heart of our discussions of distributive equality.

To capture the difference between direct and indirect theories as far as responsibility is concerned, we need to look at the notion of responsibility in a bit more

detail. To this end, recall Scanlon's (1998) distinction between *responsibility as attributability* and *substantive responsibility*. "To say that a person is responsible, in this sense, for a given action", says Scanlon explaining the former notion, "is only to say that it is appropriate to take it as a basis of moral appraisal of that person" (p 248). As opposed to that, judgments of substantive responsibility "express substantive claims about what people are required (or (…) not required) to do for each other" (p 248). The two notions are distinct; Scanlon denies that we may always infer substantive responsibility from the fact of attribution. The notion of responsibility we are interested in for questions of distributive justice is substantive responsibility, where it will be an additional question whether that notion is also tied to a notion of responsibility as attributability. In the welfarist and resourcist cases just discussed judgments of substantive responsibility are tied to judgments of responsibility as attributability, and the latter is flashed out in terms of voluntary choice.

More generally, on the direct model, substantive responsibility can enter in two ways. First, it can enter as a defeating condition: if individuals freely choose to engage in activities resulting in a shortfall of the designated good, that shortfall can legitimately be ascribed to them. It no longer represents an injustice, but "a matter of individual responsibility;" a matter of choice, instead of circumstance. Here, the individual is substantively responsible for the shortfall; the apparent deviation from equality is excused in virtue of its connection with voluntary choice. Conversely, if an individual can argue that she does not have responsibility for some shortfall, she may have a claim to be compensated. An individual with an unchosen handicap, for instance, may have a claim to additional resources, given the fact that she is not responsible for her condition.

In both cases, considerations of responsibility motivate why an individual would get more or less than others. The second way of integrating substantive responsibility into a direct theory is by means of the very concept of an equal share itself. Imagine, for example, that individuals are given an adequate supply of cash, which some rash individuals choose to burn. We might resist their claims for compensation not immediately through the concept of choice and circumstance, but simply by noting that they have indeed been given the equal share required; their foolish choice does not negate the fact that we provided what was morally demanded. Here, responsibility does not enter in a way that justifies *deviations* from the standard of equal shares; it enters, instead, by way of modifying what *counts as* an equal share to begin with.

In both cases, however, substantive responsibility is integrated into an account of distributive justice by way of qualifying the direct inference from equality of persons to the equality of their distributive shares (or, if the respective direct account is one that postulates distributive equality axiomatically, by way of modifying what precisely is postulated). Again, in both cases it will, logically speaking, be an additional question of whether judgments of substantive responsibility are tied to responsibility as attributability. In the welfarist and resourcist case discussed above, this was the case, and responsibility as attributability was understood in terms of voluntariness. This is typical of common versions of direct accounts of distributive justice. What is crucial now is that these are the only two ways in which substantive responsibility could be integrated into a direct account of distributive justice.[23]

---

[23] Within different theories, of course, substantive responsibility will be understood in different ways. Dworkin (2000), for instance, holds people responsible only for choices that derive from preferences with which they identify. Cohen (1989) and Arneson (1989)

As opposed to this – and this is the central point of difference between direct and indirect models, as far as responsibility is concerned -- the indirect model can make room for substantive responsibility within a theory of distributive justice in a markedly different way, one that not only entirely disconnects judgments of substantive responsibility from responsibility as attributability, but also utilizes the structure of indirect theories. (To be clear: the intimate connection between substantive responsibility and responsibility as attributability is a typical feature of direct theories, which tend to spell out attributability in terms of voluntariness in the manner of the welfarist and recourcist considered earlier; this connection is, however, not logically required for direct theories. The logically crucial difference stems from the fact that indirect theorists can introduce a notion of substantive responsibility that makes use of the structure of their theories.) The indirect model does not license the inference to distributive equally unless the relevant individuals stand in certain normatively important relationships. Indirect theorists can make room for substantive responsibility by way of fleshing out those relationships, and therefore they can obtain a notion of substantive responsibility that is neither available nor will seem plausible to direct theorists. A paradigmatic relationship of that sort is that of being jointly involved in a system of fair cooperation over time, a system whose members are free and equal citizens. Substantive responsibility can be integrated into this kind of indirect model by way of developing this idea of a system of fair cooperation among free and equal citizens.

---

locate the cut differently, advocating that individuals be held responsible for choices they have made in awareness of their consequences. And Roemer (1996) thinks that people should be held responsible only for those choices that similarly situated others would have avoided.

This is very abstract, but how concerns of responsibility can be integrated into indirect models is best fleshed out within that indirect theory with which we are primarily concerned, that is, Rawls's theory. In section 3 we will first develop the notion of responsibility within Rawls's theory (which will make clear what we mean by saying that "indirect theorists can make room for substantive responsibility by way of fleshing out those normatively important relationships"), and then examine two of the most sophisticated critiques of Rawls by those who charge him with an inadequate notion of responsibility, namely, Richard Arneson and John Roemer. We claim that both are committed to a direct theory of distributive equality, as well as to a version of this stance that ties responsibility as attributability to substantive responsibility. Therefore Rawls, as an indirect theorist who disengages substantive responsibility from responsibility as attributability, is not susceptible to their criticisms.

Rawls's indirect theory offers resources to develop a distinctly political notion of substantive responsibility that has little to do with the moral notion of responsibility as attributability, and that is neither open to nor will be otherwise plausible to direct theorists. While there may be good reasons for adopting a direct model of distributive justice, one cannot argue for such a model merely by showing that the indirect model does not fit with our moral account of responsibility. The indirect model does not even try to fit such a notion, since it must develop a distinct account of responsibility appropriate to the development of principles sufficient to ensure and justify the ongoing project of social cooperation.

## 3. "A Social Division of Responsibility"

3.1 A discussion of responsibility is notoriously absent from Rawls's *Theory*. Yet while responsibility never gains much prominence on his agenda, Rawls does offer the following account in several subsequent essays (emphasis added):

> This conception [of justice] includes what we may call a *social division of responsibility*: society, the citizens as a collective body, accepts the responsibility for maintaining the equal basic liberties and fair equality of opportunity, and for providing a fair share of the other primary goods for everyone within this framework, while citizens (as individuals) and associations accept the responsibility for revising and adjusting their ends and aspirations in view of the all-purpose means they can expect, given their present and foreseeable situation.[24]

Rawls claims that this understanding of responsibility is *implicit* in the choice of primary goods. That claim is puzzling. One may have thought that the issue of what "currency" to choose for distributive justice (with e.g., welfare, resources, and capabilities competing with primary goods) is both conceptually prior to and independent of how responsibility is divided between individuals and society. We will have taken a big step towards understanding Rawls's account of responsibility and its place within his theory once we understand his claim that this account is implicit in the choice of primary goods.

---

[24] This quote is from "Social Unity and Primary Goods," Rawls (1999b), p 371; see also "Fairness to Goodness," Rawls (1999b), p 284, "Reply to Alexander and Musgrave," Rawls (1999b), p 241, "A Kantian Conception of Equality," Rawls (1999b), p 261; see also "Justice as Fairness: Political Not Metaphysical," Rawls (1999b), p 407. Two questions that arise here is to what extent Rawls's theory might have *changed* between the view presented in *Theory* and his subsequent writings, and, by way of approaching that question, whether the "social division of responsibility" that appears in these later writings fits completely comfortably with everything that Rawls says in *Theory*. We will, however, not pursue these points here any further.

Interestingly, we get to the same question when studying Scanlon's (1998) discussion of Rawls's account. Scanlon asks why the parties in the original position can reasonably make their choice of principles of justice in terms of *expectations*. Reminding the reader that, for Rawls, "the choice of primary goods as a measure of expectations amounts to a certain 'social division of responsibility' between social institutions (…) and individuals," Scanlon explains:

> The idea is this: The 'basic structure' of society is its legal, political, and economic framework, the function of which is to define positions to which different powers and economic rewards are attached. If a basic structure does this in an acceptable way – if citizens have no reasonable complaint about their access to various positions within this framework or to the package of rights, liberties, and opportunities for economic reward that particular positions present them with – then that structure is just. It is up to individuals, operating within this framework, to choose their own ends and make use of the given opportunities and resources to pursue those ends as best they can. How successful, unsuccessful, happy or unhappy they are as a result is their own responsibility. (p 244)

What is puzzling is this. Plausibly it is the choice of the principles of justice that makes for a society that assures individuals of a basic stock of primary goods. Yet the principles can be chosen only *after* Scanlon's question has been settled, namely, why the parties can reasonably make their decisions in terms of *expectations*. Only after it is decided that primary goods are the "currency" *and* after parties in the original position have started to think of their choice in terms of expectations of such goods can they select principles. Scanlon can talk the way he does only if the choice of primary goods already

25

amounts to a "social division of responsibility." For then his answer to the question he raises does not depend on the availability of the principles of justice at that conceptually earlier stage. But then again we must ask: what does it mean that the choice of primary goods *amounts to* this division?

3.2 To answer this question, we must revisit the core idea of Rawls's theory. "The most fundamental idea in this conception of justice," so Rawls (2001) tells us, "is the idea of society as a fair system of social cooperation over time" (p 5). The participants of this system are considered free and equal, each with his or her own life to live, but also living together. They are regarded as equal in the sense that they are capable of engaging in social cooperation over a complete life as one among equal citizens. They are considered free in the sense that they regard both themselves and each other as having a conception of the good, and as entitled to making claims on their political institutions to be in a position to advance their respective conception of the good (provided those conceptions are in accordance with the conception of justice). Such a system of cooperation aims to realize appropriate conceptions of freedom and equality simultaneously.

Such a system can exist over time only if, first, the possession of certain goods is regulated in such a manner that the fairness of the interaction and the participants' status as free and equal citizens is maintained over time, and this leads to social primary goods as the metric of distributive justice: those are the sorts of goods participants need to this end. (This regulation will entail in particular that each citizen has an appropriate share of these goods, goods that individuals "need as citizens," Rawls (2001), section 17.) Second, such a system can exist over time only if it is the responsibility of citizens *as a collective*

*body* to assume responsibility for organizing the basic structure in such a way as to support such a system of cooperation. Otherwise, the aggregative effect of individual decisions may, over time, undermine the status of some participants as free and equal citizens. (Think of certain forms of capitalist free markets without noticeable redistributive efforts: the workings of the market may bring it about over time that some participants are no longer free and equal.) Third, cooperation can remain fair, and individuals free and equal, only if they do not, beyond reasonable limits, burden others with the costs of their decisions. Individuals must take responsibility for their ends before the background of a collective responsibility for the maintenance of the justice of the basic structure. As Rawls (1993) explains:

> [V]ariations in preferences and tastes are seen as our own responsibility. (…) [T]hat we can take responsibility for our ends is part of what free citizens may expect of one another. Taking responsibility for our tastes and preferences, whether or not they have arisen from our actual choices, is a special case of that responsibility. As citizens with realized moral powers, this is something we must learn to deal with. (p 185)

That is, individuals will have to take substantive responsibility for their tastes and preferences *regardless* of whether these tastes and preferences can also be attributed to them and are thus their responsibility also in that sense – or that is, will have to do so before the background of a basic structure whose social positions are regulated by Rawls's two principles. Responsibility as attributability, on this account, does not bear on substantive responsibility. As Scheffler (2003) puts the point:

People are asked to accept responsibility for their ends, in Rawls's sense, not because the metaphysics of the will makes it fitting that people should bear the costs of their choices, but rather because it is reasonable to expect people to make do with their fair shares. (p 27/8)

The second and third point (in terms of the enumeration in the second paragraph of this subsection) deliver Rawls's division of responsibility. This account is *implicit* in the choice of primary goods in the sense that both the former and the latter are essential to the conception of society as a fair system of cooperation among free and equal citizens over time. What accounts for the choice of primary goods as "currency" is precisely what also accounts for the endorsement of a social division of responsibility, namely, the concern to maintain such a system of fair cooperation, which in turn depends on both those components. For Rawls, then, contrary to what seemed plausible above, the question of the "currency" of justice does *not* arise conceptually prior to or independently of the question of who is responsible for what. Both questions are answered simultaneously, by way of explicating the system of fair cooperation at the core of Rawls's theory. This makes also clear why for Rawls *social* justice does not reduce to *distributive* justice conceived narrowly in terms of who should have what.[25]

---

[25] This subsection has progressed without mentioning the original position, but the latter can easily be integrated into this account. The original position is a device of representation by which we may determine what must be the case for a political society to be justified in its use of coercion. Individuals are precluded from knowing their class, plan of life, or genetic endowment; the principles derived under conditions of ignorance are deemed justified by the absence of any information which might represent illegitimate bases of argumentation. The original position is intended, through what sources of reasons it excludes, to model those forms of reason we think could be accepted as reasonable bases of agreement between free and equal persons. It is supposed to model what reasons we can legitimately provide in the justification of coercion. On such an indirect model, the ordinary moral notion of responsibility will not directly enter into an

28

Compare this outlook with Cohen's (1989) claim that "the primary egalitarian impulse is to extinguish the influence on distribution of both exploitation and brute luck," where "[a] person is *exploited* when unfair advantage is taken of him, and he suffers from (bad) *brute luck* when his bad luck is not the result of a gamble or risk which he could have avoided" (p 908). While concerns about exploitation and brute luck do have their place within the conception of responsibility developed here, they do so derivatively of the general idea of maintaining a society as a fair system of cooperation, rather than by way of capturing a "primary egalitarian impulse."

3.3 Thus tied to Rawls's notion of a fair system of cooperation, his notion of responsibility is designed to govern the assignment of burdens within such a system. As more information about society becomes available, this conception of responsibility also guides the choice of a constitution, then the design of laws, and, ultimately, the outcomes of adjudication. All of these steps are constrained by the demand that society remain a fair system of cooperation among free and equal citizens. Individuals take responsibility for their actions within the framework thus determined, which means they benefit from or assume the burdens of their choices within those constraints.

This account of Rawls's social division of responsibility should explain our claim that indirect theorists can make room for substantive responsibility by way of fleshing out the normatively important relationships whose presence they require to license inferences to distributive equality. For Rawls, these relationships are those of jointly participating in

---

account of justice. We must ask, instead, about what forms of individual responsibility are relevant from the standpoint of the justification of political coercion.

a fair system of cooperation, and it is features of that system that allow him to make room for a notion of responsibility. This notion of responsibility does not satisfy the demands on such a notion that arise in moral philosophy,[26] and it does not gain its respectability through being embeddable into an overall plausible account of the free-will problem; it is, in particular, not reducible to notions of responsibility current in that context, such as responsibility drawing on causal involvement and responsibility drawing on voluntary choice. The Rawlsian notion of responsibility cannot play such roles, as much as these moral conceptions are unsuitable to ensure that burdens are distributed in a manner that guarantees the continuation of fair cooperation among free and equal citizens.

The fact that primary goods are chosen because they are what persons need as free and equal citizens entails that moral responsibility should not be required and not be able to influence an individual's guaranteed share and thus her substantive responsibility. Just as the sort of respect individuals are due according to the Kantian moral law is not the sort of respect they must earn or of which they must show themselves worthy, so in the Rawlsian system, individuals receive shares of primary goods without needing to show themselves worthy or deserving of it. There is no room for a full-fledged notion of desert to enter Rawls's theory in a way that influences individuals' distributive shares, nor is there such room for a notion of responsibility as attributability. There is no need for the addition of such notions either.[27]


### 4. Criticizing Rawls on Responsibility, One: Arneson

---

[26] This point is also emphasized in Schaller (1997).

[27] Cf. Rawls (2001), section 20, on desert.

4.1 Richard Arneson has been concerned to argue, in a variety of contexts, that Rawls's principles of justice are inadequately sensitive to choice. Over the last decade and a half, he has defended welfarist accounts of egalitarian justice against Rawls and Dworkin. These efforts were accompanied by an exploration of ways in which the notion of responsibility could be integrated into a theory of distributive justice. In his publications around 1990 Arneson held a desire-satisfaction account of welfare (cf. Arneson (1989), (1990a), (1990b)). This account entered into his theory of equality of opportunity for welfare, which also brought to bear his ideas about responsibility on distributive justice (cf. Arneson (1989), (1991)). Arneson's more recent writings endorse an objective-list account of welfare, rejecting equality of opportunity for what he calls *responsibility-catering prioritarianism* (inspired by Parfit (1998) and related work; cf. Arneson (1999a), (2000b), and (2001)).[28] Arneson's extensive criticism of Rawls draws on both his welfarist commitments and his concern to find an appropriate place for responsibility within an egalitarian account of justice, and this concern is our subject now.

While it has taken on different forms over the years, one recurring critique is that Rawls's difference principle misallocates social primary goods by ignoring the morally relevant distinction between choice and circumstance. In particular, Rawls's vision of distributive justice illegitimately penalizes people for disabilities or innate preferences not properly understood as aspects of choice; it will treat such unchosen aspects of the person as if they were chosen. According to Arneson, this demonstrates that Rawls's

---

[28] Prioritarians are unified by demanding special consideration for those who are least well-off, according to the given metric.

vision of responsibility is inappropriate for a theory of distributive justice. We will discuss this criticism first, and then, in 4.4, turn to a second one.

The particular version of the objection we consider is set within the context of a discussion of employment (cf. Arneson (1990b)). Arneson explores whether a program guaranteeing access to productive employment could replace a program guaranteeing direct access to money and wealth. He notes that a program focusing on productive employment may illegitimately discriminate against those who have an innate aversion to wage labor and employment: among the class of persons who are both disadvantaged and unemployed, some will have a pronounced taste for leisure, so that their welfare prospects would not be enhanced by egalitarian provision of employment opportunities. These persons can be labeled *needy bohemians*. From a welfare-egalitarian standpoint, the policy of offering aid to the disadvantaged in the form of job opportunities rather than cash transfers is a good idea only if the desirable effect of screening out non-needy bohemians (who should not be targeted by welfare programs) from the policy outweighs the undesirable effect of screening out needy bohemians (who should be).

4.2 This discussion does not by itself pose an objection to Rawls. It does, however, represent a clear difference between Rawls and Arneson as regards the ways in which responsibility enters into discussions of distributive justice. Arneson's welfare-egalitarian contention is that needy bohemians, who have an unchosen aversion to waged labor ought not to be penalized for their antipathy towards work: the fact that they would be screened out by a measure that offers state-guaranteed employment instead of financial aid speaks against such a measure. To put this point in Scanlon's terms: needy bohemians

should not be held substantively responsible for their apathy towards work because this apathy is not attributable to them.

As we saw in section 3, for Rawls judgments of attributability do not bear on substantive judgments. The difference principle, which governs the allocation of primary goods to productive social and political positions, does not even attempt to distinguish needy bohemians from those who freely choose to avoid work. Having an aversion to work is insufficient to justify compensation; even if it is inborn, and thus part of an individual's set of background circumstances. On Rawls's account, individuals must adjust their plans of life to those that can be justified through the use of practical reason, or else bear the costs herself.[29] On Arneson's welfare-egalitarian account, this gets it backwards: society must adjust itself so as to ensure that unchosen preferences do not penalize individuals in the distribution of goods. Arneson, that is, ties responsibility as attributability to substantive responsibility.[30]

Arneson's account is an instantiation of the direct model of distributive justice. Where there is an inequality in what goods individuals possess, we have – on a direct account such as Arneson's – good reason to think that an objectionable inequality exists. (The "goods" here are individual welfare levels.) The notion of responsibility, on this

---

[29] See Rawls (1999b), p 368.

[30] Arneson (1997) makes a similar point: "A person who is very talented and possesses desirable traits such as charm and gregariousness may have a decided and steady preference for leisure over money-making activity, and may adopt a plan of life that involves voluntary avoidance of such activity. Even though his bank account wealth and income are low, he is living well, but Rawlsian justice lumps him together with the desperately poor who are barely able to find marginal employment and scrape by" (Arneson (1997), p 4). Later Arneson (1997) summarizes his point by saying that "[t]he difference principle mixes together deep and shallow inequalities promiscuously" (p 5).

account, enters as a defeating condition: what *prima facie* seemed like an injustice turns out not to be one once an appropriate appeal to responsibility is made. By the same token, since such a defeating condition does not apply to Arneson's needy bohemian, she *cannot* be held responsible for her plight –and a plausible theory of distributive justice must ensure that such an individual is not disadvantaged by her condition.

4.3 This is a coherent vision of distributive justice. The difficulty we seek to address is when the notion of responsibility animating this vision is taken to defeat the distributive conclusions of an indirect theorist such as Rawls. To take Arneson's needy bohemian as an objection to Rawls is to wrench a direct notion of responsibility out of context, and use it against a theory over which it has no purchase.

Arneson contends that the needy bohemian deserves compensation – deserves, that is, to be treated in the distributive realm no worse than individuals whose unchosen preferences lead to productive employment. Yet according to indirect theorists, we must ask not simply whether the individual is responsible for her preferences, but whether she can use these preferences legitimately to press claims against others. The standpoint of an indirect theorist such as Rawls is that we must determine principles by which a constitutional democracy could justify its political coercion to all those who face it. In determining whether such principles mandate treating the needy bohemian in the manner recommended by Arneson, we must ask whether there is anything in this construction demanding such treatment. It seems, however, that there is no such demand. On a treatment such as Rawls's, the desire to spend one's life in an unproductive pursuit – to avoid employment benefiting the society as a whole, as part of a system of social

cooperation – is like a demand to free ride on the efforts of others. It represents an attempt to gain the benefits of cooperation without providing any benefits.[31]

Rawls's notion of a social division of responsibility supports the idea that some disadvantages remain the responsibility of individuals. Individuals maintain the freedom to evaluate and pursue a plan of life, but within the range of such plans compatible with the cooperative system advantaging all members of society. Individuals whose natural disposition is geared towards productive labor will find this requirement easy to meet. Individuals who do not have this disposition, like the needy bohemians, bear the burden of adjusting their plans to those that might be justified to others through the process of public reason. Yet this inequality in costs does not have to be ameliorated by society.[32]

---

[31] One may say that what counts as free-riding will vary depending upon social norms; we might encounter norms permitting certain individuals to receive compensation while avoiding market-based labor. This is correct; what counts as free-riding is not a descriptive question, but a normative one, and we must admit that the precise contours of such norms may admit of more than one answer. Our point, however, persists; the mere fact of having an unchosen preference against productive labor does not, on an indirect theory, mandate compensation. The question of compensation looks not to what is chosen, but to what can be justified to others. That the latter question is likely to prove complex does not mean that it is not, for an indirect theorist, the correct one to ask.

[32] This explains, further, why Arneson could not rephrase his notion of equality of opportunity for welfare as something which might be chosen behind the veil of ignorance and, therefore, as an indirect theory. If a given division of welfare is depicted as something we might choose together – rather than a direct implication of the moral nature of welfare - then it is no longer the case that we can simply use the moral notion of responsibility to work out what distributive shares are deserved. We must, instead, seek a distinctively political reason for determining which sorts of distribution will lead to political demands. It is conceivable that Arneson's notion of equality of opportunity for welfare might be the end result of some such political process; but, in this case, the reasons for valuing the distribution would be quite different than Arneson's own. To transform Arneson's egalitarianism from direct theory to indirect theory, in other words, would be to undercut his arguments against Rawls, which rely upon a non-political notion of moral responsibility. We are grateful to an anonymous reviewer for this Journal for pressing this objection against our position.

It is the responsibility of individuals to adjust their preferences so that their participation within that system does not become the functional equivalent of a decision to defect.[33]

4.4 What we have discussed so far is Arneson's criticism that Rawls's difference principle misallocates social primary goods by not giving proper weight to responsibility. However, Arneson's work contains a second line of criticizing Rawls's treatment of responsibility. According to this objection, to the extent that Rawls does integrate responsibility (an extent that is, according to the first objection, too small), he does so in an implausible and possibly inconsistent way. Arneson objects to Rawls's "social division of responsibility" on two grounds. On the one hand this account is hard to combine with the absence of responsibility from the difference principle: if individuals are responsible for their ends, why does Rawls not hold them responsible for being among the least advantaged (cf. Arneson (1990a), (1997))? Here, so Arneson suggests, Rawls comes at least close to presenting an inconsistent theory. On the other hand, Arneson argues that this idea is too demanding and thus implausible. His endorsement of

---

[33] It is important to note, here, that Rawls's position and Arneson's may in practice lead to similar results. Both will insist that individuals who can choose to alter their plans of life to fit the requirements of cooperation will have to do so. Rawls might, further, allow some cases of truly pathological aversion to work to count as a disability, and consider such people as falling outside the set of persons understood as functioning members of society; this part of Rawls, we should note, is underdeveloped and in need of fleshing out – see Nussbaum (2006) for an analysis of these issues. The primary difference here is not with such rare and extreme cases, but with the more ordinary cases, in which the costs of altering one's plans of life are significant but not debilitating. Arneson would regard a superable but significant aversion to work as a significant moral factor, one which is deserving of compensation; Rawls, in contrast, would regard such an aversion as simply a matter of bad luck – something one must overcome in order to stand in mutually justifiable political relationships with others. Here, again, the difference is one of philosophical methodology, rather than political conclusion. We are grateful to an editor of this Journal for pressing us to be more specific on this point.

individuals' responsibility for their ends commits Rawls to "individualism with a vengeance," depriving his theory of any possibility to consider "predictably blighted lives of some of its least fortunate members" (Arneson (1997), p 16) as being within its scope.

Let us begin with the inconsistency version. An inconsistency arises only if one thinks about Rawls's division of responsibility as if it involved a notion of responsibility drawing on voluntariness, that is, a notion of responsibility as attributability. For then Rawls would have to explain why voluntariness is decisive for the assignment of costs and benefits in some contexts, but not in others. The mere insistence that, in the one case we talk about the basic structure and in the other about actions occurring within society may not satisfy the critic. For once it is assumed that what individuals have claims to is determined by their own features ("what they have chosen voluntarily") it may seem bizarre that the domain in which their choices occur should make a difference.

Yet this perspective is alien to the core of justice as fairness. The point of justice as fairness is to distribute burdens in such a manner that a fair system of cooperation among free and equal citizens is preserved, which requires a division of responsibility as suggested by Rawls. No inconsistency arises, and one would think so only if one endorses a direct theory of distributive equality. Rawls's theory, in other words, does not begin with the idea of moral responsibility, and then adapt it to politics; he begins, instead, with a political notion of what can be justified to free and equal citizens, and then develops a conception of responsibility specifically designed to work within this political theory of justice. The more ordinary notion of moral responsibility may be relevant in this process, but it need not always be so; we may not, therefore, infer any moral

inconsistency when Rawls's political notion of responsibility runs afoul of the moral notion animating Arneson's own conception.

What about the implausibility version of Arneson's objection? Is it appropriate, within a theory of justice, to hold people responsible for ends, or does Rawls not thereby assign responsibility a function it cannot plausibly have? One way of questioning whether individuals should be held responsible for their ends appears in Arneson (1997):

> [T]he distinction between inequalities arising from choice and inequalities arising from unchosen circumstances turns out to be confused, because unchosen circumstances include each individual's talent endowment, and among one's talents are talents to make and implement good choices in formulating a conception of the good and devising a plan of life. (p 14)

So by holding individuals responsible for their ends we indulge in the illusion that there is a "genuine choice" unaffected by factors that may deprive them of that responsibility. Yet since there is no such choice, this view amounts to individualism with a vengeance.

The response must be, once more, that the problem arises only if one thinks about the Rawlsian conception of responsibility as if it were a notion of responsibility drawing on voluntariness. If what individuals have legitimate claims to depends on their voluntary choices, it is easy to see how this would amount to individualism with a vengeance. That notion of responsibility, however, is Arneson's, not Rawls's, and he and other responsibility-catering egalitarians have been taken to task for it in Anderson (1999). The Rawlsian system, to use a formulation due to Ripstein (1999), is based on the idea that "equality and responsibility need to be understood together, in light of the idea of

38

reciprocity" (p 21).[34] Before this background, "individualism with a vengeance" must be seen as a phenomenon rather alien to Rawls's theory. Such vengeance is a threat to responsibility-catering egalitarianism that uses a choice-based notion of responsibility.

4.5 There are some difficulties in these answers to Arneson.  The first is the distinction between disability and expensive preference.  Individuals with disabilities preventing them from participation in productive employment are guaranteed compensation and resource shares on Rawls's account.  There is a relevant difference between those who *cannot* work, and those who *prefer* not to work.[35]  Nonetheless, it must be recognized that this line may be difficult to draw in particular cases.  Rawls assumes that individuals will be able to adjust their preferences to productive pursuits, while the disabled will be unable to adjust their circumstances to permit employment.  In the majority of cases, this assumption will be correct.  The distinction, nonetheless, might prove problematic.  A further difficulty is found in the notion of productive employment itself.  What counts as cooperation – what counts as productive employment – is not an empirical question, but a normative question.  As such, it admits of different answers.[36]  Whether the competitive market is an adequate guide to those forms of pursuit providing benefits to others is a

---

[34]  This is actually what Ripstein (1999) says is the "central theme" of the book.

[35] Goodin (1985) insists that, on theories making use of the idea of society as a cooperative venture for mutual advantage, those too weak to participate productively must rationally be excluded from membership in the community.  As applied to Rawls, this seems to us inaccurate; members of political community who are unable to work because of unchosen disability are nonetheless entitled to have the political coercion they face justified to them – a process of justification which may involve a model such as that developed by Rawls.  See Goodin (1985) p. 156.

[36] Cf. Waring (1990).

question we cannot address here.  Nevertheless, such a question must be answered for indirect theories such as Rawls's to be acceptable.[37]


## 5. Criticizing Rawls on Responsibility, Two: Roemer

5.1 Like Arneson, John Roemer has developed a series of criticisms of Rawls's concept of responsibility.  The one we examine deals with questions of effort and desert. Roemer takes his arguments (in conjunction with objections to Dworkin and Harsanyi) to entail that veil-of-ignorance arguments do not support egalitarian conclusions, advising egalitarians not to use the veil-of-ignorance devise. We have nothing to say about Roemer's arguments against Dworkin and Harsanyi, but argue that Rawlsians can maintain the connection between the veil of ignorance and egalitarian conclusions. The

---

[37] What about Arneson's (1997) point that "[t]he difference principle mixes together deep and shallow inequalities promiscuously" (p 5), where "deep" inequalities are beyond the individuals' control (and thus not their responsibility), whereas "shallow" inequalities are due to voluntary choice (and thus are)? Curiously, Rawls also talks about "deep inequalities," insisting that it is such inequalities to which his principles of justice apply in the first instance (Rawls (1999a), p 7). What Rawls counts as deep inequalities are inequalities arising from unequal starting positions. Regulating such inequalities is the target of Rawls's theory. "Shallow" inequalities then are those not due to unequal starting points, but, presumably, due to the actions of individuals in a manner that cannot be account for by a reduction to unequal starting points. Rawls's distinction between deep and shallow inequalities is at odds with Arneson's, which is drawn in terms of the voluntariness of actions: "shallow" inequalities are those due to voluntary actions and require no redress, whereas "deep" in inequalities are due to circumstances not brought about by voluntary actions and thus do. The Rawlsian distinction between inequalities arising because of the socio-economic constitution of society and inequalities arising because of individual actions within such structures is ill-captured in terms of voluntary and involuntary actions. For if inequalities arise because of unequal staring points, they do so regardless of whether they are maintained by voluntary or involuntary actions. The language of voluntariness ignores that concern. The domain of Rawlsian principles of justice are *institutions*, rather than individual actions. Institutions, unlike actions, are ill-assessed in terms of the distinction between the voluntary and the involuntary. As far as principles of justice for the basic structure are concerned, that distinction does little work.

work will be done, once again, by the distinction between direct and indirect accounts of distributive equality.

Where Arneson contests that Rawls illegitimately punishes people for unchosen unproductive plans of life, Roemer is concerned that Rawls's refusal to regard effort as giving rise to desert fails to give people proper credit for freely chosen decisions to work hard. This contention is expressed by a comparison between effort and plans of life:

> [Rawls] assumes that individuals do not know their life plans behind the veil of ignorance, although we may suppose they know the joint distributions of such life plans and all other resources. But if life plans are not morally arbitrary, then why should individuals in the morally correct posture for making the social contract not know them? And why should they not know their propensities to expend effort, or, rather, that part of effort for which individuals are morally responsible? After all, the veil of ignorance is only supposed to shield individuals from knowledge of their morally arbitrary features. (Roemer (1996), 175-176)

This critique is best understood in terms of Rawls's notion of responsibility. Both plans of life and knowledge about effort are hidden in the original position. Yet why should we hide from individuals the knowledge about their propensity for effort, where that effort is the result of free choice? Rawls seems guilty of falsely ascribing to circumstance what should be understood as free choice, thereby refusing to credit individuals for choices for which they actually are responsible. Roemer insists that individuals should know their life plans behind the veil. But if so, says he, many will reject the difference principle. So Rawls must drop either the claim that the veil admits all

and only morally relevant information (depriving it of its purpose), or he else his claim that the original position offers an impartial viewpoint delivering egalitarian conclusions.

5.2 This critique, like Arneson's, rests on an assumption that the relevant notion of responsibility is one found within the direct model of distributive equality. It is an attempt to read the direct model back into Rawls's original position. Whatever is morally arbitrary, on this reading of Rawls, cannot serve to ground a difference in the allocation of goods. The critique charges that Rawls is inconsistent in his handling of plans of life and effort; if both are morally arbitrary, neither can justify inequalities in distributive shares. Since effort at least partly represents free choice, Rawls errs by refusing to give individuals credit for these choices. Respecting individuals, for Roemer, means compensating them for choices for which they are morally understood as responsible.

Yet the original position is not designed to eliminate all aspects of the person for which we cannot claim responsibility. It is not a replication or formalization of the ordinary notion of moral responsibility. It is, instead, used to generate principles that could legitimately govern cooperation. As Rawls says,

> [w]e must specify a point of view from which a fair agreement between free and equal persons can be reached; but this point of view must be removed from and not distorted by the particular features and circumstances of the existing basic structure. (Rawls (2001), p 15)

He goes on to explain:

> In the original position, the parties are not allowed to know the social positions or the particular comprehensive doctrines of the persons they represent. They also

42

do not know persons' race and ethnic group, sex, or various native endowments such as strength and intelligence, all within the normal range. We express these limits on information figuratively by saying the parties are behind a veil of ignorance" (Rawls (2001), p 15; cf. Rawls (1999a), sec. 24; Rawls (1993), p 24f). Crucially, the language of morally arbitrary features is ill-suited to depict the purpose of the original position. For the original position serves to capture conditions under which an agreement about principles governing a system of fair cooperation can be reached. It is not simply a way of spelling out the moral notion of responsibility for choice; it is, instead, a method for working out principles for fair political agreement – and the notion of responsibility it implies is internal to this task, rather than something derived from our moral general ideas of moral responsibility.

One's plan of life, on this analysis, is not best described through the idea of moral arbitrariness. It has a distinct place within the process of generating legitimate cooperative principles. It shares with effort the fact that plans of life, by themselves, cannot justify differential shares within a cooperative system; I cannot simply cite the inherently superior nature of my way of life in arguing for a higher share of distributive goods. The original position therefore excludes knowledge of plans of life to ensure impartiality in the selection of principles. Such plans are not arbitrary, but they cannot be legitimately appealed to in the construction of legitimate principles we can expect contracting parties to endorse. A plan of life, however, is distinct from effort. Plans of life, from within the methodology espoused by Rawls, must be respected to the extent that they can be made compatible with other plans of life, as part of the broader process of respecting individuals. Given that the cooperation we seek is to occur within the

43

context of a broad diversity of plans of life and views about the good, we must respect these plans of life if we are to determine principles all participants have reason to accept.

We may thus examine plans of life with reference to the idea of political coercion. We seek to justify such coercion to free and equal persons. We cannot insist that a single way of life is inherently superior to another. Nor can we make our own way of life an illegitimate burden upon others, by refusing to cooperate in productive activities underlying social cooperation. The veil of ignorance, understood in this light, does not model conclusions about moral arbitrariness, but a political conception of what burdens we can legitimately place upon others. Plans of life are excluded in the original position not because they are morally arbitrary, but because of the unique place of such plans in the justification of political coercion. Rawlsian liberal egalitarians have, therefore, no need to abandon the use of the veil of ignorance in their theoretical explorations.

5.3 What, then, of effort? While effort is similarly excluded from what is known in the original position, the reason for its exclusion is distinct. Both effort and plans of life, we may acknowledge, are morally mixed. They are both the result of a combination of free choice and unchosen circumstance. From the standpoint of a theory concerned to justify political coercion to free and equal persons, the question arises: can such an aspect of persons legitimately ground an inequality in distributive shares? The answer to such a question may be controversial, but to answer *no* seems at least a plausible response. Principles allowing even partly caused aspects of persons to influence distributive shares directly in this way might be legitimately resisted by reasonable persons.

It is worth noticing that the question asked by Roemer depends upon an ability to discount circumstance from choice that seems utterly unavailable to limited creatures such as ourselves. The only question we may ask is whether effort, taken as a whole, may legitimately influence distributive shares; focusing upon that proportion of effort due to choice assumes knowledge we do not, and could not, have. This is why Rawls terms reward for effort *impractible*, rather than illegitimate. Were we able to perform the hypothetical calculation imagined here, we would perhaps have reason to do so. Given the very human nature of our knowledge, however, we cannot; and given this fact, distributive principles rejecting rewards for effort are in line with notions of responsibility that can be used in indirect models. It is also worth noticing, finally, that the question of effort is perhaps less important than is generally assumed. The only form of effort for which we receive no credit is effort *within* a given profession or social position. Cooperative systems such as Rawls's, however, allow differential rewards to flow to different positions; and they allow talents and ambitions to determine the allocation of persons to positions.[38] The talent set of any given individual, however, is far from static. An individual's effort in the development of new talents, the refinement of existing talents, and the acquisition of new skills enables her to attain new positions.

All this, of course, is perfectly in harmony with Rawls's system, which does not seek to equalize individual talents but marshal those talents for the benefit of all. What this means, however, is that individual choice to expend effort will often be rewarded with additional distributive shares. Individuals expend differential amounts of efforts

---

[38] Rawls's system may also allow differentials within positions, when those differentials result from choices for leisure time versus other goods. A dentist who chooses to treat more patients may legitimately gain more income than one who chooses to treat fewer. We are grateful to an anonymous reviewer for raising this point.

within a given social position, and they expend differential amounts of efforts in professional and social advancement. The former sort of differential gives individuals no claim upon others for additional goods; the latter, however, may legitimate differential rewards, as individuals develop differential sets of talents and skills. Rawls is frequently accused of insufficient sensitivity to effort by those who assume talents to be static. When this view is replaced with a more dynamic conception of talent, we find that individual effort may be relevant to distributive shares in a real, albeit indirect, manner.

In a nutshell, then, Roemer, like Arneson, does not have a case against Rawls, and the reason for this is that he, again like Arneson, endorses a direct model of distributive equality, rather than an indirect one.


## 6. Conclusions

Nothing we have said here mandates the choice of an indirect model of distributive equality. We hope only to have demonstrated that the distinction between indirect and direct models of distributive equality is a valid, and important, distinction. We hope, further, to have demonstrated that the conceptions of responsibility accompanying these models are similarly distinct. Those theorists who would seek to critique the validity of an indirect theory of distributive equality, therefore, have more work to do than simply to demonstrate the ways in which such theories are incompatible with a direct notion of responsibility. The indirect model is immune from such criticisms, inasmuch as it relies upon a distinct notion of responsibility involved in the justification of political coercion, rather that the perhaps more immediate moral notion of moral responsibility. The critics of indirect theorists such as Rawls may, in the end, be correct; but such critique cannot

proceed in the manner in which it has generally been undertaken. A fresh start is required; and, until this task is undertaken, indirect theorists such as Rawls may take their theories to have survived the challenge posed by responsibility-catering egalitarians.

## Literature

Anderson, Elizabeth. 1999. "What is the Point of Equality?" *Ethics* 109: 287-337

Arneson, R. 1989. "Equality and Equal Opportunity for Welfare." *Philosophical Studies* 56: 79-95

Arneson, Richard. 1990a. "Primary Goods Reconsidered." *Nous* 24: 429-54

Arneson, Richard. 1990b. "Is Work Special? Justice and the Distribution of Employment." *American Political Science Review* 84: 1127-1147

Arneson, Richard. 1990c. "Liberalism, Distributive Subjectivism, and Equal Opportunity for Welfare," *Philosophy and Public Affairs* 19: 158-194.

Arneson, Richard. 1997. "Rawls, Responsibility, and Distributive Justice." Forthcoming in *Justice, Political Liberalism, and Utilitarianism: Themes from Rawls and Harsanyi*. Ed. M. Salles and J. Weymark. Cambridge: Cambridge University Press

Arneson, Richard. 1999. "Against Rawlsian Equality of Opportunity." *Philosophical Studies* 70: 77-112

Arneson, Richard. 2000. "Welfare Should be the Currency of Justice." *Canadian Journal of Philosophy* 30: 497-524

Arneson, Richard. 2001. "Luck and Equality." *Proceedings of the Aristotelian Society*, supp. vol. (2001), pp. 73-90

Barry, Brian. 1995. *Justice as Impartiality*. Oxford: Oxford University Press

Beitz, Charles. 1979. *Political Theory and International Relations*. Princeton: Princeton University Press

Blake, Michael. 2001. "Distributive Justice, State Coercion, and Autonomy." *Philosophy and Public Affairs* 30: 257-297

Cohen, G. A. 1989. "On the Currency of Egalitarian Justice." *Ethics* 99: 906-44

Dworkin, Ronald. 1981. "What is Equality? Part 2: Equality of Resources." *Philosophy and Public Affairs* 10: 283-345

Dworkin, Ronald. 2000. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge: Harvard University Press

Goodin, Robert. 1985. *Protecting the Vulnerable: A Reanalysis of Our Social Responsibilities*. Chicago: University of Chicago Press

Kymlicka, Will. 2002 (2nd ed.). *Contemporary Political Philosophy*. Oxford: Oxford University Press

Lippert-Rasmussen, Kaspar. 2001. "Egalitarianism, Option Luck, and Responsibility." *Ethics* 111: 548-579

Nagel, Thomas. "The Problem of Global Justice," 33 *Philosophy and Public Affairs* (2005) 113-147.

Nussbaum, Martha. 2006. *Frontiers of Justice: Disability, Nationality, and Species Membership.* Cambridge: Harvard University Press.

Pauer-Studer, Herlinde. 2002. *Constructions of Practical Reason: Interviews on Moral and Political Philosophy.* Palo Alto: Stanford University Press

Pogge, Thomas. 1989. *Realizing Rawls*. Ithaca: Cornell University Press

Rawls, John. 1993. *Political Liberalism.* New York: Columbia University Press

Rawls, John. 1999a. *A Theory of Justice*. Revised Edition. Cambridge: Harvard University Press

Rawls, John. 1999b. *Collected Papers*. Ed. S. Freeman. Cambridge: Harvard University Press

Rawls, John. 1999c. *The Law of Peoples*. Cambridge: Harvard University Press

Rawls, John. 2001. *Justice as Fairness. A Restatement*. Ed. E. Kelly. Cambridge: Harvard University Press

Ripstein, Arthur. 1994. "Equality, Luck, and Responsibility." *Philosophy and Public Affairs* 23: 3-23

Ripstein, Arthur. 1999. *Equality, Responsibility, and the Law*. Cambridge: Cambridge University Press

Risse, Mathias. 2005. "What Do We Owe to the Global Poor." *Journal of Ethics* 9: 81-117

Roemer, John. 1996. *Theories of Distributive Justice*. Cambridge: Harvard University Press

Roemer, John. 2001. "Three Egalitarian Views and American Law." *Law and Philosophy* 20: 433-460

Roemer, John. 2002. "Egalitarianism Against the Veil of Ignorance." *Journal of Philosophy* Vol. XCIX, No. 4: 167-184

Scanlon, T. M. 1998. *What We Owe To Each Other*. Cambridge: Harvard University Press

Schaller, Walter. 1997. Expensive Preferences and the Priority of the Right: A Critique of Welfare-Egalitarianism. *Journal of Political Philosophy* 5: 254-273

Scheffler, Samuel. 2003. "What is Egalitarianism?" *Philosophy and Pubic Affairs* 31: 5-40

Singer, Peter. 1972. "Famine, Affluence, and Morality." *Philosophy and Public Affairs* 1: 229-243

Singer, Peter. 1993. *Practical Ethics.* Cambridge: Cambridge University Press

Singer, Peter. *One World.* Princeton: Princeton University Press, 2005.

Waring, Marilyn. 1990. *If Women Counted: A Feminist Economics*. Sand Francisco: Harper

Wolff, Jonathan. 2002. "Addressing Disadvantage and the Human Good", *Journal of Applied Philosophy*: 207-218