

## **Egalitarianism and Responsibility**

(published in *Journal of Ethics* 3, No. 3 (1999))

Richard J. Arneson                      May, 1999

The problem of social justice can arise in the absence of social interaction. This point emerges directly from an important passage in John Rawls's *A Theory of Justice* where he is arguing for an opposed conclusion. Rawls argues that the primary subject of justice is the basic structure of society, the way that major social institutions work together to "determine the division of advantages from social cooperation." He writes,

"The basic structure is the primary subject of justice because its effects are so profound and present from the start. The intuitive notion here is that this structure contains various social positions and that men born into different positions have different expectations of life determined, in part, by the political system as well as by economic and social circumstances. In this way the institutions of society favor certain starting places over others. These are especially deep inequalities. Not only are they pervasive, but they affect men's initial chances in life; yet they cannot possibly be justified by an appeal to the notions of merit or desert."<sup>1</sup>

This passage contrasts deep and shallow social inequalities and associates deep inequalities with the basic structure of society. Deep social inequalities are inequalities in people's initial life prospects that are imposed on them in ways that are entirely beyond their power to control. We might think here of inequalities in life prospects among children born into different places in the social hierarchy of wealth and status. In contrast, shallow inequalities might be conceived as ones that arise among people who are equal in life chances initially but then choose to behave in ways that render them differentially meritorious or deserving and that also render them unequal in subsequent

life prospects, perhaps in ways that are justifiable given their different choices for which they can reasonably be deemed responsible. (The categories of deep and shallow inequalities so understood do not exhaust the possibilities.) Rawls writes as though deep inequalities are brought into being by the basic structure of society, but it does not seem that this is the whole story. Individuals are born with different talents and traits that affect their initial life prospects independently of the social structure, so that someone born with great native talent into the very bottom of the social hierarchy might yet have extremely favorable initial life prospects.

In fact we can imagine deep inequalities arising entirely outside of any social structure in a Robinson Crusoeesque setting. Imagine a number of persons, each living on a separate island. The islands contain different qualities and amounts of soil and other resources. The individuals are endowed with different qualities and amounts of personal traits such strength, physical coordination, and intelligence that enable them to satisfy goals they might adopt.<sup>2</sup> The islands are situated so that social interaction and cooperation between individuals on two different islands is impossible, but we suppose that somehow the individuals have accurate knowledge of one another's life conditions. Finally, suppose that it happens to be the case that one-way traffic between some of the islands is possible: from some islands it is possible to place resources in a boat that will drift with the tides to the shores of another island.

If the existence of deep inequalities triggers a problem of social justice, then this brief account provides all that is needed to raise this problem. The inequalities in the internal and external resource holdings of the island inhabitants are given in the circumstances that fall on each agent quite independently of any choice or action anyone

takes for which one might be deemed to be personally responsible. A theory of justice gives guidance as to when deep inequalities are justified and requires us to alter the situation if present conditions are unjustified. Here I suppose that justice requirements will give rise to moral obligations that trump other considerations and that are legitimately enforceable. The imaginary island scenario convinces me that there can be social justice obligations in the absence of society, if society is conceived as a mutually advantageous scheme of cooperation, a set of institutions and practices.

Robert Nozick invokes this same island example to illustrate his claims (1) that no one has any social justice obligation to aid anyone else when each lives separately and (2) that if there are no obligations to aid in the absence of society, the introduction of voluntary trade and other schemes of social cooperation does not give rise to obligations to aid the needy.<sup>3</sup> I accept claim (2). If no credible distributive justice principles imply that individuals are under obligations of justice to aid the disadvantaged in the absence of society, it is hard to see how complicating the picture by introducing social cooperation changes this result. (Interesting deontological conceptions of equality challenge this claim, but I set them aside in this essay.<sup>4</sup>)

My main aim in this essay is to explore possible rationales for the position that in the absence of society justice requires that fortunate individuals should give up resources to improve the life prospects of those whose initial conditions are unpropitious. Call this the generic egalitarian intuition. An act-utilitarian will hold that if and only if transferring resources one controls to a person whose resource endowment is poor maximizes the sum of utility (well-being), then one should make the transfer. I shall stipulate that to qualify as egalitarian, a social justice principle must recommend transfers

from well off to badly off persons in circumstances when the resource transfer is not utility-maximizing. What varieties of egalitarianism can be identified, and which, if any, are plausible?

#### THE "SIMPLEST BASIS."

The essence of the account of the islanders might seem to be that some of them face worse life prospects than others through no fault of their own. If there are obligations of justice to rectify this situation, the simplest basis for these obligations would seem to be to assert that (1) it is morally bad--unjust and unfair--if some people are worse off than others through no fault of their own. To infer from Principle (1) that anyone is obligated to transfer aid one might add to it the claims (a) if one can eliminate something that is morally bad at reasonable cost to oneself and without giving rise to anything comparably morally bad, one should do so and (b) the antecedent clause of (a) is satisfied for some islanders. Our concern here will be to explore alternative characterizations of what is morally bad in the situation of the island inhabitants. This initial "simplest basis" is not so simple, for Principle (1) is ambiguous. Although Larry Temkin has urged Principle (1) as a canonical formulation of egalitarianism, in fact (1) does not unequivocally affirm any sort of egalitarian principle.<sup>5</sup>

Notice first that Principle (1) is fully compatible with Principle (2): It is morally bad—unjust and unfair—if some people are as well off as others through no merit of their own. Taken together, Principles (1) and (2) are compatible either with a qualified affirmation of equality or (modulo a qualification to be noted) with a principle of moral meritocracy, Principle (3): It is morally good—just and fair—that each person be exactly

as well off, by comparison with others, as she deserves to be.<sup>6</sup> The more deserving one is, the better off one should be.

Principle (3) would have it that people who are equally deserving should be equally well off, but the goodness of equality of distribution here is entirely a byproduct of what really matters, namely, that well-being should vary with desert. Principle (1), conjoined with Principles (2) and (3), arguably does not call for the individuals with more resources to ship some of their resources, if they can, to the island inhabitants who are less well endowed. Instead whether redistribution is called for depends on the extent to which individuals turn out to be deserving or undeserving. A wait and see attitude would be the appropriate response to initial inequality of fortune, at least if one supposes that what matters is not a correspondence between good fortune and deservingness at any one moment of one's life, but rather correspondence over one's life as a whole. Seen this way, an initial undeserved inequality of fortune is morally inconsequential provided that over the long run of each person's life, the extent of one's good fortune matches one's deservingness. Hence if we are seeking a principled explanation for the judgment that better off islanders do have obligations to aid their less fortunate neighbors in the original scenario, moral meritocracy looks to be a poor candidate for the job.

This conclusion might be resisted on the ground that virtually any principle to which one might appeal to justify a transfer of resources from better off to worse off at a single moment might withhold any recommendation to transfer if the unit of significance for the application of the principle is each individual's life as a whole. In this regard moral meritocracy is no different from the rival principles we shall pit against it.

There is something right about the initial claim, however. It is not that according to moral meritocracy there is some initial reason to equalize the distribution of good fortune, a reason that might be outweighed by subsequent events. There is no initial reason to equalize at all. It is as though just before the start of a horse race one observed that there is some reason to award every horse the prize for finishing first, since none has so far run faster or slower. But none has run at all, so the principle of distribution according to merit is silent.

One might have a derived deservingness reason to alter the initial distribution of resources among the island inhabitants, but this reason does not per se provide any reason for distribution from better off to worse off. One might hold that an initial distribution should give all persons a fair opportunity to show themselves deserving. A poor initial resource holding might dampen an individual's propensity to behave meritoriously, so that one judges that others who are better endowed initially enjoy a more favorable opportunity to reveal their disposition to be meritorious in their conduct. But some individuals may be most likely to behave meritoriously when their circumstances are disadvantageous. In the face of adversity, they tend to shine, but they are likely to be corrupted by initial good fortune. Hence if one wished to conjoin to the moral meritocracy ideal the norm that people should be given a fair opportunity to behave meritoriously, that double norm of meritocracy would not establish an initial presumption in favor of transferring resources from better endowed to worse endowed, but rather a presumption in favor of transferring resources in whatever direction would establish in the circumstances a fair opportunity for all to show themselves deserving.

Another possible construal of (1) is that it is intended to affirm the value of equal distribution on the condition that everyone is equally deserving. On this construal, equality of distribution is morally valuable for its own sake, but not unconditionally, just on the condition that everyone is equally deserving. Notice that this construal of Principle (1) is not so very different from my initial construal, because (1) interpreted as affirming the conditional value of equality is itself compatible with both (2) and (3).

The proposal that it is good that people are equally well off on the condition that they are equally deserving cries out for further clarification. Consider two possible worlds: in world A, everyone is equally deserving and equally well off. In world B, the aggregate amount of deservingness is the same as in world A, the number of persons inhabiting each of the worlds is the same, but in world B people are unequally deserving and unequally well off, and the more deserving one is, the better off one is. Question: Does someone who affirms that equality is good on the condition that all who are equally well off are equally deserving thereby commit herself to a preference for world A over world B? I suppose the answer is NO. Affirming that equality is good if a certain condition is satisfied, one thereby leaves it open whether it is good, bad, or indifferent that the condition be satisfied.

The task of clarifying Principle (1) is not yet complete. A puzzling feature of (1) is that read literally, it is only concerned with inequalities that are caused in particular ways. “It is bad if some are worse off than others through no fault of their own” says that inequalities are bad unless caused by the fault of those who end up on the short end of the inequality. Suppose that Simon Legree is very undeserving; he is a thoroughly rotten person. Suppose that Joan of Arc is a thoroughly virtuous person, very deserving by any

plausible standard of deservingness. Suppose also that Simon Legree leads an unenviable life that is low in well-being, and Joan of Arc leads an enviable existence rich in well-being. But the good fortune that Joan gets is just that—good fortune. Her good fortune is not causally linked to her virtue. She happens to be incredibly lucky at some critical juncture of her life, and her happiness and fulfillment are entirely traceable to her good luck. In contrast, Simon Legree was vicious, but his vice happened to be irrelevant to the genesis of the poor quality of his life. He was just incredibly unlucky at key junctures of his life. His bad fortune is the mirror image of Joan's good fortune. But notice that the inequality in the quality of the lives had by Joan and Simon is condemned by Principle (1). As it happens, Simon is worse off than others through no fault of his own.

I would urge that if someone believes that the more deserving should have better lives and the less deserving should have less good lives, it is not plausible in addition to insist that correspondence between good fortune and degree of deservingness is only morally desirable when the two are causally linked, the individuals' meritorious (unmeritorious) conduct causing her good (bad) fortune. This leaves open the further question whether a causally linked correspondence between people's deservingness and the good or bad fortune they enjoy is morally better, other things being equal, than an otherwise similar correspondence between people's deservingness and their good fortune that lacks this causal linkage. Be that as it may, the moral meritocracy view constituted by Principles (1) and (2) goes astray by failing to register the idea that it is good that there be correspondence even in the absence of causal linkage. (Principle (3) alone does register this idea and to this extent better expresses the moral meritocracy ideal.)



The proposed seemingly most simple basis for urging the moral desirability of transfers from better endowed to worse endowed individuals in the island scenario has turned out not to be any sort of basis for that position. To explicate the moral case for transfer, we need to explore further.

#### EQUALITY.

Perhaps a simpler thought than Principle (1) is needed, such as Simple Equality: Other things being equal, it is morally bad—unfair and unjust—if some are worse off than others.

The moral ideal of Simple Equality is controversial. Some find it attractive. Others are repelled. Our concern is whether equality is intrinsically morally valuable, valuable for its own sake, rather than as a possible means to other moral goals. One objection against equality, if proposed as valuable for its own sake, is that its appeal comes about through a halo effect, when it is confused with other values that are worthwhile for their own sakes. One possible deceptive halo effect is that equality might be affirmed on the ground that it is morally bad if some are worse off than others through no fault of their own, but it then turns out that it is also affirmed that it is morally bad if some are as well off as others through no merit of their own, and that the desirability of correspondence between good fortune and one's merits is the true source of these judgments, equality being only a byproduct of what is deemed intrinsically morally valuable.

Harry Frankfurt rejects economic egalitarianism, the ideal that all persons should have the same income and wealth, for three reasons.<sup>7</sup> One refers to another halo effect. He asserts that when advocates of equality bemoan an unequal distribution, what they

actually are finding objectionable is not the fact that some have more than others, but rather that in the circumstances they envisage, those who have less are leading lives that are grim and impoverished. But if inequality between the affluent and the wealthy is not morally troubling, then if inequality between the destitute and those better off is morally offensive, what is bad is not merely the inequality but the suggested badness of the lives of the destitute coupled with the hint that this badness is avoidable by means of redistribution.

Frankfurt's second point is that a concern for equality per se is fetishistic, because a concern for how one's condition compares to the situation of others is a concern that is tangential to concerns that matter. One's life goes better or worse in virtue of its intrinsic qualities, not in virtue of the relationship in which those qualities stand to the corresponding qualities of the lives of others. What does matter—this is Frankfurt's third point—is not whether one has as much as others, but whether one has enough. This doctrine of sufficiency holds that an individual has enough resources when he is contented with his holdings and does not take an active interest in gaining more. One might reasonably be contented with one's present holdings either because one's life is going well or because although one's life is not going well its defects would not likely be ameliorated as a result of gaining more resources.<sup>8</sup>

Frankfurt directs his arguments specifically at equality of wealth and income, but they apply more broadly against any conception of equality of resources, and arguably against an egalitarianism of well-being as well. If it makes sense to be contented in such a way that one does not actively seek more resources, it also makes sense to be contented in such a way that one does not actively seek a greater degree of the satisfactions one

might gain from the use of those resources. When one has enough goal satisfaction, according to Frankfurt's view, one regards one's level of satisfaction as acceptable, and is not taking an interest in getting more.

Frankfurt's second and third arguments against equality prove unconvincing under examination. At least, that is what I shall try to show. The first argument requires the acknowledgement that something other than equality is intrinsically morally important, but does not strictly force the denial that equality is not intrinsically morally important. But despite these doubts about Frankfurt's arguments, I find plausible the bare claim that equality and more generally, how one person's condition compares to that of others, do not matter morally intrinsically. When I contemplate cases of levelling down, in which equality can be achieved by destroying the advantages now possessed by better-off persons without in any way improving the condition of anyone else, I do not judge that there is one respect in which the outcome of levelling down is an improvement—it creates equality—even though perhaps, all things considered, the change is not morally desirable. The levelling down seems a waste, pure and simple, and everyone's having the same does not seem in any way intrinsically worthwhile. But I have no argument to support this response, and others may disagree.

Frankfurt's concern that the moral demand for equality is fetishistic is hard to credit. If a cake is to be divided between two people, and each wants all of the cake for himself, and neither of the individuals has any special claim to an extra share of cake, the fair solution may be to divide the cake equally. The fact that neither of the two potential recipients of the cake cares a hoot about equality does not rule out the possibility that equal division is the morally required solution. Nor is equality's candidacy for the rule of

standard of fair division impugned by the supposition that one's life goes better or worse in virtue of its intrinsic features, not in virtue of how those features compare to features of the lives of others. (It seems to me that my life could go better or worse depending on whether my ambition to write the world's greatest poem is fulfilled or not, but let that pass.) The norm of equality, like the norm of moral meritocracy, identifies fairness with how one person's condition compares with the condition of others. Frankfurt might yet be correct to assert that a concern for equality is fetishistic, but this is simply a fancy way of saying that equality is not per se morally important and does not provide a reason for accepting that assertion.

One reason to be skeptical about the probative force of Frankfurt's claim that concern for equality is fetishistic is that even if we concede that what matters from a prudential standard is how well one's life goes, as assessed according to its intrinsic features, not how it compares with how others' lives are going, this concession leaves it entirely open that from the standpoint of morality, equality and more generally how one person's lot compares to that of others might matter fundamentally, for its own sake. Frankfurt points out if one thinks about what kind of life one wants for one's children, and what kinds of considerations one wants one's children to regard as intrinsically important, how their resource share compares with that of others will not figure in these wants. But to envisage a parental concern for one's own child is to envisage a frame of mind in which concern for the child's own welfare naturally predominates and impartial moral considerations take a back seat. So this sort of example does not tend to show that distributive equality is not morally important.

Another potential source of confusion here is that Frankfurt might be right that equality in the distribution of income resources is not intrinsically morally important, not because equality does not matter, but because the distribution of resources is not of intrinsic, just instrumental significance. I find plausible the idea that the distribution of material stuff to people matters morally not for its own sake but for what people are enabled to do and be by means of this stuff. But this thought does not have any tendency to show that equality in the distribution of whatever ultimately matters morally should not be intrinsically morally significant.

Frankfurt's doctrine of sufficiency is hard to evaluate, because it is hard to disentangle it from the idea that satisficing, aiming to achieve some threshold level of benefit that is deemed satisfactory, is in many contexts a sensible strategy as judged by a maximizing standard. This can be so when seeking further gains above the satisfactory level involves the risk of wasted effort for marginal gain and even a risk of falling below the satisfactory level in the end. Regarded not as means to maximization but as in themselves desirable, satisficing and the doctrine of sufficiency run afoul of the simple point that many valuable, choiceworthy goods do not have an upper bound. If one unit of pleasure is good, two is better, and if there were a way for a human to experience a superhuman level of pleasure of a billion units or more, that would be better still. The construction of a good novel is good, but it would be better to construct a better one. The doctrine of sufficiency implies that if an individual has a choice of two life strategies, which are in relevant respects identical except that the reasonably expected benefit from choosing the first is greater than the reasonably expected benefit of choosing the second,

there is nothing irrational about choosing the second, provided that its expected benefit level is good enough. Incoherence looms.

There is an easy way to accept Frankfurt's first point against the moral value of equality without being forced all the way to the dubious sufficiency doctrine. For simplicity, consider Frankfurt's criticism as it applies to equality of utility. If the plight of a miserable person can be alleviated by transferring resources to her from a moderately happy person, and an equally large gap in utility levels between an extremely happy person and an ecstatically blissful person could be remedied by a similar transfer of resources, the norm of equality would have it that exactly as strong a reason supports each of the proposed transfers. Frankfurt notes that this is counterintuitive. The first transfer is morally imperative, the second is morally inconsequential, which shows that concern for equality is not what is driving our response to these cases.

It does not seem to me that the reason in favor of transfer disappears as we shift from the first case to the second. It just lessens. This could be so because equality matters, but is not the only thing that matters. The extra thing that seems to move us is the imperative of eliminating dire need or terrible life circumstances. One way to explain this pattern of response is to suppose that there is some threshold of decent existence, and we should give priority to boosting individuals up to this threshold. This would explain why achieving a transfer from a well-off person to a person in miserable grinding poverty is morally more urgent than achieving a comparable transfer from someone who is super-rich to someone who is merely rich.

Another way to explain the response that equalizing when those with the short end of the inequality are in grim straits is morally more important than equalizing when

those with the short end of the inequality are already quite well off abandons the claim that equality is per se important and instead appeals to the prioritarian family of principles. Prioritarianism holds that institutions and practices should be set and actions should be chosen to maximize moral value, with the moral value of achieving a benefit (avoiding a loss) for a person being greater, the greater the size of the benefit, and greater, the lower the person's lifetime expectation of benefit prior to this gain. I assume that benefit levels should be measured in terms of well-being (utility, prudential good). Prioritarianism is an attractive version of egalitarian justice, but further discussion of the threshold approach to distributive justice will lay the groundwork for the exploration and defense of prioritarianism.

#### THE THRESHOLD OF A DECENT QUALITY OF LIFE

Discussing the doctrine of sufficiency, Frankfurt appeals to an idea that resembles it but is perhaps more plausible. The idea is that when inequality of resources among persons appears bad, what is undesirable is not the inequality per se, but rather the fact that we are imagining a scenario in which the worse off are leading worse lives than anyone should have to endure. Following up this thought, we might identify a minimally decent standard of living, a quality of life below which no individual should be forced. The associated imperative of distributive justice would be to minimize the number of people who suffer a below-threshold existence (here I leave aside the issues that arise when policy affects population size, and consider only cases in which population is fixed.).

Notice that depending on the quality of the lives in the scenario of the individuals each living isolated on a separate island, and depending on the level at which the

threshold of decent existence is set, the norm of minimizing below-threshold lives could just as well recommend transfers of resources from worse off to better off islanders as in the reverse direction. If each islander is now below the threshold, but transfers from worse off to better off can push more individuals to the threshold than could any alternative, then “Minimize below threshold existence!” recommends these transfers.

I have two doubts about minimizing the number of people who must endure below-threshold lives, if this norm is proposed as fundamental principle of justice. One doubt is whether it is possible to specify a nonarbitrary level of minimal decency. A related doubt is why any such level should have special importance in deciding what we owe to one another. Suppose that some people are leading lives of hellish quality. If resources are transferred to them their lives will be improved significantly, but it is not feasible to institute transfers that would boost their life prospects up to the minimally decent level. For the sake of the example we can suppose further that the improvements in the lives of the hell inhabitants that transfers to them would achieve would increase aggregate human utility to a greater level than any alternative policies we could institute. Nevertheless the norm of minimizing below threshold lives recommends transferring resources from the hell inhabitants, making their lives significantly worse, if transfers from them to some better off persons would boost those better off persons up to the level of minimal decency. This result is an indictment of the norm that gives priority to increasing the number of those who lead a minimally decent existence.

The idea that it is morally a more urgent concern to bring people up to some threshold decent quality of life than to improve people’s conditions of life above the threshold can be interpreted differently, so that my second objection does not apply. The



decent threshold norm might hold that achieving utility gains and avoiding losses for people below the threshold count for more in determine what should be done than achieving gains and avoiding loses for those above the threshold. If the extra weight assigned to below-threshold gains is uniform, I would object that it is morally more important to get a one unit gain in utility (the pleasure of a cold drink of water, for example) to someone in hellish conditions, far below the threshold, than to obtain the same gain for someone in almost adequate life circumstances, just below the threshold. The same point applies to above-threshold gains. If one amends the decent threshold norm so that utility gains count for more, the lower in well-being the beneficiary of the gain prior to its receipt, we are close to a prioritarian view. The only remaining objection is the moral arbitrariness of setting the decent quality of life threshold at any particular point. But if the threshold does not make a difference to the evaluation of policies and actions, its arbitrariness would be trivial.

### PRIORITARIANISM

Improvements in the decent threshold view render it into prioritarianism. Recall that according to prioritarianism, the moral value of gaining a benefit (avoiding a loss) of a given size for a person is greater, the lower the person's level of benefits (measured on an absolute rather than a comparative scale) prior to receipt of the benefit (avoidance of the loss).<sup>9</sup> A principle that calls for maximizing benefits will recommend transfer of resources from better off to worse off persons provided that the total of benefits is thereby increased. A principle that calls for maximization of moral value as computed by prioritarianism will recommend transfer of resources in the island scenario even if net benefits are lessened, either because the resources are damaged in transfer or because the

givers of the resources are better at converting resources into benefits than the recipients, provided that any such lessening of benefits is outweighed by the moral value of provision of benefit to those who are badly off.

According to prioritarianism, equal distribution is sometimes valuable as a means to what matters, but has no value in itself. Variants of the island scenario in which prioritarianism would not recommend equal distribution reveal prioritarianism to be attractive in its policy implications. If the island inhabitants are teetering near the edge of subsistence, so that transfers of resources that would result in equal distribution would result in swift death for all, prioritarianism does not attach any moral value to equal distribution, but will prefer the distribution that brings as many island inhabitants as possible over the threshold of assured survival to better times. If transferred resources mostly rot in transit before reaching their intended destination, at some rate of rot prioritarianism will cease to recommend transfers which now would produce no net moral benefit even though they would continue to render the distribution of benefits among islanders more nearly equal. On principle, prioritarianism could recommend transfers from worse off island inhabitants to better off inhabitants, if a small sacrifice by the worse off would generate a sufficiently large benefit increase for the better off to offset the prioritarian discount of improvements in benefit levels for the already better off. (Imagine that seeds that will not sprout in the soil on the islands of the worse off [we assume that only one type of transfer between islands, transfer of seed, is feasible] and that are hardly useful at all to the worse off could grow into glorious blossoms on the islands of the better off.)

“Prioritarianism” names a type of position, not a specific principle. To get a specific principle one needs a function that determines, for each increment on an absolute benefit scale, the moral value of securing a small benefit for a person at that benefit level. At one end of the prioritarian scale one gets virtually no weighting and an identification of maximal moral value with maximization of benefits; at the other end of the scale one gets leximin. When I contrast prioritarianism with other views I have in mind not the generic position but a restricted family of priority weightings in the middle of the range, but I have nothing useful to say about how to identify a specific principle.

#### MORAL MERITOCRACY

Prioritarianism as stated attracts the objection that it wrongly denies any fundamental place in principles of distributive justice to considerations of personal responsibility and deservingness.

Prioritarianism can and no doubt should incorporate responsibility and deservingness in an instrumental role. For example, we can imagine setting laws and social rules so that individuals are held responsible for actions and their outcomes in the sense that they will be subject to praise or blame, reward or punishment, depending on their quality or in the sense that the individual will be required to pay the costs of her actions that fall on others or to absorb the costs of her actions that fall on herself under specified conditions. Some such laws and rules might be part of the best package of laws and rules that could be put in place in order to achieve the prioritarian moral goal. Notions of deservingness, virtue, and vice might be incorporated in the means to achieve prioritarianism in a similar way. But norms of responsibility and deservingness so regarded are not intrinsically morally important, important for their own sake, not merely

as a means to some other moral goal. That an agent who behaves responsibly gets a better deal in life than one who behaves irresponsibly does not render a situation morally more desirable in itself, apart from possible desirable gains to weighted well-being that achieving a better deal for the responsible might promote. The issues that arise in this context are familiar from discussions of utilitarianism, to which prioritarianism is structurally similar.

The issue arises in the world of isolated islands we have imagined. Suppose that two islanders happen to have equivalent external and internal resources as they begin their adult lives. A behaves virtuously and B does not; since there is no social interaction, let's suppose we can identify being virtuous with being prudent. We might wonder whether A has been endowed with an initially unnoticed personal resource, a capacity for prudent choice-making and choice-executing, which B lacks. But this is not the case. Upon examination it is revealed, so far as we can discern, that it is no more difficult or painful for B to make and execute prudent choices than it is for A. On an appropriately fine-grained account of personal responsibility, A behaves far more responsibly, far more virtuously, than B, and the result is that A now enjoys an excellent quality of life while B's quality of life is poor. Now an earthquake causes the tides to shift and C could send resources either to A's or B's island that would improve their life prospects to the same degree. This allocation by C will be in the nature of a lump-sum allocation, a windfall that lands on A or B and will not affect their future incentives to behave in ways that would achieve the prioritarian moral goal to a better or worse degree. C can simply make A or B better off, and that is an end to the story; no further consequences are relevant. Since both can be helped to the same extent at the same cost

to C and since B is far worse off, prioritarianism grinds out the recommendation that C morally ought to give aid to B not A. The example just illustrates the point that responsibility and deservingness can only be instrumentally valued by the prioritarian, and since here no instrumental gains are in play, responsibility and deservingness cannot be valued at all.

In the absence of some hard determinist or other demonstration that it would make no sense to incorporate notions of responsibility and deservingness as determinants of what fundamentally matters in moral principle, the demotion of these norms to secondary status is suspect.

The prioritarian could stand her ground and insist that instrumental responsibility is responsibility enough. I want to mention the possibility of a more accommodating response. One might simply accept that priority and personal responsibility are both of intrinsic moral significance and seek a principle that appropriately registers both values. One is then shopping for the best version of responsibility-catering prioritarianism. For example, one might amend prioritarianism as follows: Institutions and practices should be set and actions chosen to maximize moral value, with the moral value of achieving a gain (avoiding a loss) for a person being (1) greater, the greater the amount of well-being for the person the gain (averted loss) involves, (2), greater, the lower the person's lifetime expectation of well-being prior to receipt of the benefit (avoidance of the loss), (3) greater, the larger the degree to which the person deserves this gain (loss avoidance). We ought to maximize well-being weighted by priority and responsibility.

A mixed view of this type, that aims to incorporate more than one, but just a few, values into the principles of distributive justice, must answer criticisms from two sides.

On the one side, one must defend the claim that the plural values one has singled out as morally fundamental really are separate and independent moral values that are morally important in their own right. On the other side, one must defend the claim that the few values one has selected, among many other candidates, should be given privileged treatment. Why pluralism at all? And if pluralism, why just a little pluralism—why not let a hundred values bloom in one’s fundamental moral principles?<sup>10</sup>

Notice that instrumental norms of responsibility and deservingness might diverge widely from the norms of responsibility and deservingness we are inclined to accept as fundamental after reflection. This divergence need not constitute incoherence. There would just be two different sets of values that would need to be integrated and balanced. A common objection against monistic utilitarianism notes that the ideas of responsibility and deservingness that would make sense as tools for utility maximization might be deeply in conflict with our common-sense judgments of these matters. In a similar way, instrumental responsibility and deservingness might sharply conflict with fundamental responsibility and deservingness. In particular, if we allow that the extent to which it is reasonable to hold people morally responsible for their bad choices depends on the extent to which the circumstances thrust on them beyond their power to control render it easy or difficult, pleasant or painful, to make good rather than bad choices, then judgments of fundamental responsibility and deservingness might be very sharply revisionary with respect to ordinary common-sense judgment. Perhaps on a suitably fine-grained account of personal responsibility, my responsibility for slight misdeeds such as being snide to my friend on some occasion might be full, so that I am very blameworthy and deserving of punishment, whereas Adolf Hitler’s and Charles Manson’s degree of responsibility for

their heinous misdeeds might be so vanishingly slight that they are hardly at all blameworthy and deserving of punishment. Exemplary severe punishment for Hitlers and Mansons and no punishment at all for Arnesonian peccadilloes might be what is justified by instrumental norms of responsibility. To see how much we care about fundamental responsibility as compared to instrumental responsibility, we should imagine possible cases like this in which their implications would sharply diverge and examine to what extent we would after reflection give allegiance to one or the other. Another example of this type of comparison: Many of us suppose that a competitive market economy conjoined to redistributive tax and transfer policies could do tolerably well to bring about the achievement of prioritarian values—the maximization of human well-being weighted by the value of priority to the worse off. But a competitive market responds to supply and demand, not fine-grained or for that matter coarse-grained estimations of different individuals' degrees of deservingness and responsibility. If we imagine institutions that would do better to bring about distribution of the good in accordance with people's true deservingness, but at significant cost of priority-weighted aggregate well-being, would we then be inclined to scrap the competitive market in order to institute a tolerably adequate moral meritocracy? Such hypothetical judgments can be used to determine how much weight should be assigned to responsibility in a responsibility-catering prioritarianism that accords responsibility values fundamental not just derivative status.

#### KAGAN'S CHALLENGE

Maybe some combination of our views that good fortune should go to the meritorious explains all of the plausible judgments about cases that we might have

thought could only be rationalized by invoking egalitarian premises. Shelly Kagan presses this interesting line of thought in his essay “Equality and Desert.”<sup>11</sup> Applied to the isolated islanders scenario, Kagan’s insight is that we can explain and rationalize what I have called the generic egalitarian intuition purely by appeal to the idea that people should get what they morally deserve without admitting any appeal to egalitarian or prioritarian values.

To elucidate this challenge to egalitarianism, let us suppose that what renders individuals differentially deserving is their virtue, which can be measured on a cardinal interpersonal scale. One individual is absolutely more deserving than another if her virtue is greater.

Kagan distinguishes comparative from noncomparative desert. Suppose that we can measure the overall amount of good fortune that individuals get and compare their good fortune to their virtue. Noncomparative desert is the idea that for each person with a given amount of virtue, there is some definite amount of good fortune corresponding to this virtue that she should get. From the standpoint of desert, it is bad if someone gets less than she deserves, and bad if she gets more than she deserves. The amount of good fortune that a person deserves, given her level of virtue or moral merit, Kagan calls her “peak,” because when the person has this amount of good fortune, the situation so far as she is concerned is as good as it can be from the standpoint of desert.

Comparative desert according to Kagan is the idea that each person’s absolute desert level (her peak) in conjunction with that of other persons, determines how well off one should be as compared to others. If I am less absolutely deserving than you, then you should be getting more good fortune than I—you are comparatively more deserving. If



we are equally absolutely deserving, we should be getting the same good fortune—we are comparatively equally deserving. From the standpoint of comparative desert, if we are both getting more than we absolutely deserve, but we are absolutely equally deserving, then comparative desert says we should get the same good fortune. If you are better off than I, and we are both better off than we absolutely deserve, comparative desert (in a two-person universe) says that I should get more good fortune, so that we are equally well off, since we are equally absolutely deserving. If two people both have less than they absolutely deserve, and one has more than the other, but are equally absolutely deserving, then it is better, from the standpoint of comparative desert, if the one who has more should be made less well off, so that both are equally well off, since they are equally comparatively deserving. The implications of comparative and noncomparative desert sometimes conflict, so a theory of moral desert would have to decide how much weight should be assigned to each in determinations of the overall goodness of situations from the standpoint of desert and in determinations of what should be done.

Alongside the notion of being absolutely deserving Kagan introduces the idea of being specifically deserving. If A is absolutely more deserving than B (i.e., A has a higher peak), but B is currently farther below her peak than A is below hers, then B is specifically more deserving than A. This means that given their current absolute desert levels and good fortune levels, it is better from the standpoint of noncomparative desert that B should get a benefit that is available than that A should get it. One is more specifically deserving of a benefit in the offing than another person if one is farther below one's peak than the other is. Kagan also notes one might hold that as someone gets farther and farther below her peak, the more and more desirable it becomes, from the

standpoint of desert, that the person gain a unit of good fortune, and similarly, as someone's good fortune level gets farther and farther beyond her peak, the less and less desirable it becomes, from the standpoint of desert, that the person gain a unit of good fortune. Kagan calls this phenomenon curved desert because of the way it would be represented if one measures an individual's good fortune on one axis and the goodness, from the standpoint of desert, of the individual's having any specified amount of good fortune, on another axis of a coordinate system.

With these ideas in place, return to the separated islanders scenario. Some islanders happen by sheer luck to be well endowed with personal talents and to live on an island that is rich with resources. Others are badly off. The egalitarian intuition is that the better off should transfer some of their resources to the worse off, compensating them for their bad fortune.

Kagan's discussion is directly concerned only with the evaluation of outcomes, not with the theory of right action and policy, what we ought to do. But he in effect notes that notions of deservingness can support the claim that the situation of the islanders as described is morally undesirable and that the situation in which a more equal distribution obtains is morally superior. Any theory of right that tells us we have some obligation to bring about better rather than worse outcomes will support the case for transfers from fortunate to unfortunate islanders in these circumstances.

If the separated islanders are equally deserving, but unequally well off, then considerations of deservingness alone, absent any concern for any egalitarian value, can yield the result that it is better that benefits go to the worse off. Consider a two-person case. First, comparative desert considerations indicate that since both are equally

absolutely deserving (they have the same peak), and one is worse off, the one who is worse off is specifically more deserving. So if we compare the status quo and an alternative feasible position in which the better off individual gives up some resources and the worse off individual gains that amount of resources, the latter position is morally more desirable according to comparative desert. The same conclusion can be supported by considerations of curved desert. If two individuals have the same peak, but one is farther below his peak than the other, then the value, from the standpoint of desert, of securing a benefit or preventing a loss for the worse off person is greater, given curved desert, than the value, from the standpoint of desert, of securing a benefit or preventing a loss for the better off person. This desert-based preference for gaining benefits for the worse off can hold true even though the net result is a loss of aggregate utility (so utilitarianism would not recommend the shift of benefits to the worse off). Pure considerations of desert with no admixture of equality can then explain the so-called generic egalitarian intuition.

This last claim requires a qualification. Prior to the establishment of any desert claims, I asserted earlier in this essay, the idea that good fortune should be made to correspond to one's level of individual desert does not support any redistribution, egalitarian or otherwise, in the separated islanders scenario. But once people act so as to render themselves variously deserving, a desert-based principle of distribution will have something to say about who should get what. Kagan's challenge to egalitarianism is the proposal that once we understand the distributive implications of the desert-based view, no egalitarianism remains plausible.

Further grounds for thinking that desert trumps equality are provided by considering cases in which desert considerations and equality considerations pull in opposed directions. Kagan considers this case. Suppose A is far more absolutely deserving than B and is also far better off (enjoys a higher level of good fortune). A is a saint and B a sinner. Although A is better off than B, A is less well off than she absolutely deserves to be (she is below her peak), whereas B, though worse off, already enjoys more good fortune than he deserves (he is beyond his peak). Equality considerations (including prioritarian evaluation) would say that it is morally better for B, the worse off, to receive a gain that is in the offing than A. But from the standpoint of desert, this is not so. Why confer a benefit on B, the sinner who already has more than he deserves, at the expense of A, the saint, who has less than she deserves? Reflecting this case, Kagan reports that he sees no moral value at all to preferring the outcome in which B, the worse off, gets an extra benefit rather than A. This is to say that at least in this important case of comparison, equality has no weight at all in competition with desert. Kagan explores other possible cases in which we might suppose that equality has some weight, and finds that desert considerations taken together suffice to explain all of our moral judgments, so he tentatively concludes “that egalitarianism should not be accepted after all.” He emphasizes that his central concern is to show that because of the complexity in the structure of the theory of desert, it is far more difficult than usually thought to construct examples that suffice to test the extent to which our moral judgments include a commitment to egalitarianism or desert.

Does acknowledgement of complexity in the theory of desert provide grounds for doubting that what I have called the generic egalitarian intuition rests on egalitarian views at all? I strongly doubt it.

My main ground for suspicion about Kagan's interesting result is that neither noncomparative desert nor comparative desert as characterized by Kagan is a plausible normative idea.

Noncomparative desert is defined in terms of the idea that there is some definite level of well-being or good fortune that a person, given her degree of virtue deserves. For each individual there is some peak beyond which further increases in her well-being level make the outcome worse rather than better from the standpoint of desert. A related but weaker idea is that for each person, given her level of attained virtue, there is some definite level of well-being or good fortune that she deserves, such that further increases in well being beyond this plateau do not make the resulting situation better from the standpoint of desert. But why suppose there is any such peak or plateau? One reason for denying this claim, which I find compelling, can be appreciated if one imagines a utopia of desert in which economic and social life is arranged so that each person gets now exactly the level of well-being that we think she absolutely deserves. Now imagine that some breakthrough technology is discovered, or we discover vast new exploitable natural resources on the moon. This discovery enables us to triple the well-being of every person on earth for many generations. The adjustment of society to the new situation might result in some people getting more than they comparatively deserve. But we can suppose that institutions and practices adjust marvelously, so that the huge gains in the available means for securing well-being are distributed so that each person gets in comparative

terms exactly what she deserves. But this scenario, involving as it does every individual on earth gaining more good fortune than she absolutely deserves, is according to noncomparative desert morally undesirable. This claim is so implausible that it discredits the noncomparative desert idea that generates it.<sup>12</sup> One might conclude from this discussion that desert claims are inherently comparative.

However, comparative desert as Kagan defines it, constructed from the idea of noncomparative desert, inherits the defects of the latter.

Comparative desert also is like the straight equality norm in that it can yield assessments of outcomes that are inconsistent with the minimally controversial Pareto norm. Suppose we can alter the status quo by making someone worse off without making anyone else better off. The Pareto norm says that this alteration from the status quo would not be an improvement. Comparative desert can recommend a change from the status quo to a Pareto-inferior position. Suppose that you and I are equally absolutely deserving, but you are now far better off than I. We cannot do anything to make me better off, but we could destroy some of your benefits, so that you are worse off than you were, and enjoy a level of well-being closer to what I have. Comparative desert assesses the situation in which your position is worsened as better from the standpoint of desert than the status quo. If we respond to the hypothetical case that making you worse off does not improve the situation in any moral respect, we are rejecting the idea of comparative desert as formulated by Kagan.

Since Kagan's examples and interpretation that are intended to show that desert considerations crowd out equality considerations and that we have no need to appeal to

the latter all rely on his notions of noncomparative and comparative desert, which are unacceptable, the argument against egalitarianism collapses.

Look again at the case Kagan considers decisive, where it was said we could help either a saint, already well off, who has far less than she deserves, or a sinner, badly off, who has more than he deserves. But if we scratch the unpalatable idea of peaks, what are left with is this description: One person is worse off than another, but the worse off person is the less deserving. We simply have two considerations that pull in opposite directions. Which consideration dominates in a given case and determines whom we ought to help if we can only aid one of the two claimants depends on the facts of the case and the relative weights our moral principles attach to the two considerations. Nothing here supports the claim that the consideration that tells us to tilt in favor of helping a person, the worse off he is prior to receipt of the benefit (the prioritarian idea) has nil or insignificant weight.

It is possible to develop ways of conceiving desert that do not employ the Kagan conceptions of comparative and noncomparative desert and do not violate the Pareto norm and that imply, for some cases, the generic egalitarian intuition. But so far as I can see the idea that deservingness should entirely displace any egalitarian norm, to the extent that it is plausible at all, rests on the idea that there is some set amount of good fortune to which one is entitled for any level of deservingness—the idea of peaks that we have seen reason to reject. One might insist that each person should ideally gain good or bad fortune proportionate to what she deserves, but I have suggested that once one begins to sort out what we can reasonably be held responsible for and what renders us truly deserving, and cleanly distinguishes these ideas of “true” deservingness from

instrumental versions of responsibility and deservingness tailored to serve other moral goals, the idea that these desert-based notions should rule the roost seems dubious.

Moreover, to the degree that responsibility and deservingness considerations do seem intrinsically morally important, they can be integrated with prioritarianism to yield responsibility-catering prioritarianism. My tentative conclusion is that prioritarian theories provide a sensible accounting of the generic intuition and are worth further exploration.

But this is a rather tentative conclusion. It turns out that disentangling the moral basis for redistribution even in very artificially simple environments without social interaction is a difficult and, so far as this essay is concerned, unfinished task.

---

<sup>1</sup> . John Rawls, A Theory of Justice (Cambridge: Harvard University Press, 1971), p. 7.

<sup>2</sup> What counts as a talent depends on the environment. The personal traits that are useful in one environment may be less so, or not useful at all, in another environment.

<sup>3</sup> . Robert Nozick, Anarchy, State, and Utopia (New York: Basic Books, 1974), p. 185. David Gauthier invokes a similar desert island case to make a similar point in Morals by Agreement (Oxford: Oxford University Press, 1986).

<sup>4</sup> . On such deontological conceptions of equality, see Thomas Nagel, “Justice and Nature,” Oxford Journal of Legal Studies, vol. 17, no. 2 (Summer, 1997) pp. 303-321.

<sup>5</sup> . Larry Temkin, Inequality (Oxford: Oxford University Press, 1993), p. 13. In a footnote on this page Temkin explains that he uses the expression “through no fault of their own” as an abbreviation for “through no fault or choice of their own.” This phrasing introduces the possibility that nonfaulty choice by an agent leading to a lowering of her position below the average level that others enjoy might be morally unproblematic. To my mind this formulation suggests an opportunity-oriented rather than a desert-based view; these are different. Temkin also formulates his position as the claim that “undeserved inequality is always objectionable” (p. 12), but this formulation contains the same ambiguity I note in the text, as it is compatible with the further claim that undeserved equality is always objectionable.



---

<sup>6</sup> . The principle of moral meritocracy comes in two flavors, comparative and noncomparative. The text states a comparative version. According to noncomparative moral meritocracy, there is an absolute scale of moral virtue or deservingness, and corresponding to each level on this scale is an appropriate degree of good fortune. Each person should get what she deserves according to this scale, regardless of what others do or get. See Joel Feinberg, "Noncomparative Justice," Philosophical Review, vol. 83, no. 3 (July, 1974), 297-338; also Shelly Kagan, "Equality and Desert," in What Do We Deserve? A Reader on Justice and Desert, ed. Louis P. Pojman and Owen McLeod (New York and Oxford: Oxford University Press, 1999, 298-314. I discuss the ambiguities in the Temkin formulation of egalitarianism in "Comments on Satz and Temkin," presented at APSA meeting, September, 1996.

<sup>7</sup> . Harry Frankfurt, "Equality as a Moral Ideal," Ethics, vol. 98, no. 1 (October, 1987) 21-43.

<sup>8</sup> . Michael Slote defends the idea that a rationally egoistic person might seek what is less than best for herself. On Slote's view, it can be rational to satisfice, rather than optimize. See his Beyond Optimizing (Cambridge and London: Harvard University Press, 1989).

<sup>9</sup> . On prioritarianism, see Derek Parfit, "Equality or Priority?", The Lindley Lecture (University of Kansas, 1995).

<sup>10</sup> . Another issue would need discussion in a thorough treatment. If one modifies prioritarianism by stipulating that achieving well-being gains for the responsible and deserving is morally important for its own sake, the question arises why one does not further modify the view, by holding that the aggregate amount of deservingness and perhaps its distribution across persons should be added to the values to be maximized. I have here been supposing that being deserving is a matter of doing the best one can with the cards that fate has dealt one. On this view, society cannot, just by reshuffling the deck and getting an individual better cards, improve the likely deservingness of the individual, since one's true deservingness score is adjusted to offset changes in circumstances. But this metaphor needs unpacking.

<sup>11</sup> . Shelly Kagan, "Equality and Desert."

<sup>12</sup> . This objection was brought to my attention by Brad Hooker.