

Possession, Desert, and Distributive Justice

Having clarified the status of Rawls' motivational assumptions, we can now put his theory of the person and his theory of justice side by side to check for the fit between them. In this way we can work within the argument from reflective equilibrium, by asking whether the theory of the person contained in the original position corresponds to the principles of justice it must both shape and reflect. Of special interest for this purpose is the difference principle, the principle that permits only those inequalities that work to the benefit of the least advantaged members of society. We shall see in this chapter that an adequate defense of the difference principle must presuppose a conception of the person unavailable on deontological assumptions, that we cannot be subjects for whom justice is primary and also be subjects for whom the difference principle is a principle of justice. A central focus will be on the role of desert in distributive justice, and on the conception of possession it requires. To explore these issues, we begin by contrasting Rawls' views with various competing distributive theories, and in particular the rival, but in some ways strikingly similar theory defended by Robert Nozick (1974).

LIBERTARIANISM TO EGALITARIANISM

From a practical political point of view, the positions of Rawls and Nozick are clearly opposed. Rawls, the welfare-state liberal, and Nozick, the libertarian conservative, define between them the clearest alternatives the American political agenda has to offer, at least where issues of distributive justice are concerned. And yet, from a philosophical point of view, they have much in common. Both define their positions in explicit opposition to utilitarianism, which each rejects on the grounds that it denies the distinction between persons. Both offer instead a rights-based ethic said to secure the liberty of individuals more completely. Although Nozick's account of rights owes much to Locke, both appeal to Kant's precept to treat

every person as an end and not merely as a means, and seek principles of justice that embody it. Both deny that there exists any social entity above or beyond the individuals who comprise it. As Nozick writes, echoing Rawls in both principle and rhetoric,

Side constraints upon action [that is, unqualified prohibitions] reflect the underlying Kantian principle that individuals are ends and not merely means Side constraints express the inviolability of other persons. But why may not one violate persons for the greater social good? Individually, we each sometimes choose to undergo some pain or sacrifice for a greater benefit or to avoid a greater harm But there is no *social entity* with a good that undergoes some sacrifice for its own good. There are only individual people, different individual people, with their own individual lives. Using one of these people for the benefit of others, uses him and benefits the others. Nothing more To use a person in this way does not sufficiently respect and take account of the fact that he is a separate person, that his is the only life he has (1974: 30-3).

The moral side constraints upon what we may do, I claim, reflect the fact of our separate existences. They reflect the fact that no moral balancing act can take place among us; there is no moral outweighing of one of our lives by others so as to lead to a greater overall social good. There is no justified sacrifice of some of us for others (1974: 33).

Both theorists emphasize what Rawls calls 'the plurality and distinctness of persons' and what Nozick calls 'the fact of our separate existences'. This is the central moral fact that utilitarianism denies and that an individualistic, rights-based ethic affirms. On this moral fact, and on the importance of rights, Rawls and Nozick emphatically agree. And yet Rawls arrives at a theory of justice on which social and economic inequalities are permitted only in so far as they benefit the least well off, while Nozick holds justice to consist in voluntary exchanges and transfers alone, ruling out redistributive policies altogether. How then do their theories of justice come so sharply to diverge? Fortunately, the point of divergence can be located with some precision, since Rawls, in developing his second principle of justice (the one containing the difference principle), lays out a line of reasoning that begins with a position similar to Nozick's and ends with his own.

Rawls considers three possible principles by which the distribution of social and economic benefits might be regulated or assessed: natural liberty (similar to Nozick's 'entitlement theory'), liberal

equality (akin to a standard meritocracy), and democratic equality (based on the difference principle). The system of natural liberty defines as just whatever distribution results from an efficient market economy in which a formal (i.e. legal) equality of opportunity prevails, such that positions are open to those with the relevant talents. Rawls finds this principle inadequate on the grounds that the distribution it sanctions tends simply to reproduce the initial distribution of talents and assets; those substantially endowed will wind up with substantial shares, and those with meager assets will end with meager results. Where the outcome tends simply to reproduce the initial distribution, it is possible to call it just only on the additional assumption that initial endowments were justly distributed. But this assumption cannot be established. 'The initial distribution of assets for any period of time is strongly influenced by natural and social contingencies', and as such is neither just nor unjust but simply arbitrary. And since there is nothing to recommend the justice of initial endowments, to install them in the name of justice is to incorporate the arbitrariness of fortune, nothing more. 'Intuitively, the most obvious injustice of the system of natural liberty is that it permits distributive shares to be improperly influenced by these factors so arbitrary from a moral point of view' (72).

The principle of liberal equality seeks to remedy the injustice of natural liberty by going beyond formal equality of opportunity and correcting, where possible, for social and cultural disadvantages. The aim is a kind of 'fair meritocracy', in which social and cultural inequalities are mitigated by equal educational opportunities, certain redistributive policies, and other social reforms. On the principle of liberal equality, the ideal is to provide all an 'equal start', so that those with similar native talents and capacities and a similar willingness to exercise them would have 'the same prospects of success regardless of their initial place in the social system, that is, irrespective of the income class into which they are born. In all sectors of society there should be roughly equal prospects of culture and achievement for everyone similarly motivated and endowed. The expectations of those with the same abilities and aspirations should not be affected by their social class' (73).

But while liberal equality represents an improvement over the system of natural liberty, 'intuitively it still appears defective'. Fair

opportunity, however full, is too weak an assault on the arbitrariness of fortune.

Even if it works to perfection in eliminating the influence of social contingencies, it still permits the distribution of wealth and income to be determined by the natural distribution of abilities and talents. Within the limits allowed by the background arrangements, distributive shares are decided by the outcome of the natural lottery; and this outcome is arbitrary from a moral perspective. There is no more reason to permit the distribution of income and wealth to be settled by the distribution of natural assets than by historical and social fortune (73-4).

Once we are struck by the arbitrariness of initial endowments determining life prospects generally, we are bound on reflection to be as disturbed by the influence of natural contingencies as social and cultural ones. 'From a moral standpoint, the two seem equally arbitrary' (75). The same reasoning that leads us to favor a 'fair meritocracy' (as in liberal equality) over a purely formal equality of opportunity (as in natural liberty) naturally leads us further to seek what Rawls calls the democratic conception. But it seems clear that the democratic conception cannot be found in a simple extension of the principle of fair opportunity. For one thing, it would be virtually impossible to extend opportunity so completely as to eradicate even those inequalities stemming from social and cultural conditions alone. The institution of the family, for one, makes it 'impossible in practice to secure equal chances of achievement and culture for those similarly endowed' (74). But even if compensatory education and other reforms could fully, or even nearly, correct for social and cultural deprivation, it is difficult if not vaguely forbidding to imagine what kind of social policies would be required to 'correct' in a comparable way for the contingencies of natural fortune. What is needed, then, is a conception that nullifies the *effect* of these differences while at the same time acknowledging their intractability.

Some commentators, and particularly those hostile to the principle of democratic equality, describe the next logical step as a move from equality of opportunity to equality of result. In their view, any theory of justice that rejects a meritocratic conception for the moral arbitrariness of its distributive consequences must necessarily be committed to a kind of levelling equality requiring constant readjustment of distributive shares to correct for persisting differences of

by no means the only democratic alternative to a meritocratic regime, nor is it the principle adopted by Rawls. The difference principle is not synonymous with equality of result, nor does it require the levelling of all differences between persons. 'It does not follow that one should eliminate these distinctions,' writes Rawls. 'There is another way to deal with them' (102). Rawls' way is not to eradicate unequal endowments but to arrange the scheme of benefits and burdens so that the least advantaged may share in the resources of the fortunate. This is the arrangement that the difference principle seeks to achieve. It defines as just only those social and economic inequalities that work to the benefit of the least advantaged members of society. Taken together with the principle of offices and positions open to all under fair equality of opportunity, the difference principle defines Rawls' conception of democratic equality.

The difference principle is not simply a fuller version of the principle of fair opportunity; it attacks the problem of arbitrariness in a fundamentally different way. Rather than transform the conditions under which I exercise my talents, the difference principle transforms the moral basis on which I claim the benefits that flow from them. No longer am I to be regarded as the sole proprietor of my assets, or privileged recipient of the advantages they bring. 'The difference principle represents, in effect, an agreement to regard the distribution of natural talents as a common asset and to share in the benefits of this distribution whatever it turns out to be.' In this way, the difference principle acknowledges the arbitrariness of fortune by asserting that I am not really the owner but merely the guardian or repository of the talents and capacities that happen to reside in me, and as such have no special moral claim on the fruits of their exercise.

Those who have been favored by nature, whoever they are, may gain from their good fortune only on terms that improve the situation of those who have lost out. The naturally advantaged are not to gain merely because they are more gifted, but only to cover the costs of training and education and for using their endowments in ways that help the less fortunate as well. No one deserves his greater natural capacity nor merits a more favorable starting place in society (101-2).

By regarding the distribution of talents and attributes as a common asset rather than as individual possessions, Rawls obviates the

of social and natural contingencies. When 'men agree to share one another's fate' (102), it matters less that their fates, individually, may vary. This is why, although the difference principle tends to 'redress the bias of contingencies in the direction of equality', it 'does not require society to try to even out handicaps as if all were expected to compete on a fair basis in the same race' (101).

Rawls acknowledges that the difference principle and in particular the notion of endowments as 'common assets' clash with traditional conceptions of individual desert. 'There is a natural inclination to object that those better situated deserve their greater advantages whether or not they are to the benefit of others' (103). Rawls' response is that this conception of individual desert is a mistake, as the general argument from arbitrariness suggests. 'It seems to be one of the fixed points of our considered judgments that no one deserves the place in the distribution of native endowments, any more than one deserves one's initial starting place in society' (104). The claim that a person deserves at least what he achieves through his own effort is more intuitively plausible, but even the willingness to strive conscientiously may largely be determined by social and natural contingencies. 'The assertion that a man deserves the superior character that enables him to make the effort to cultivate his abilities is equally problematic; for his character depends in large part upon fortunate family and social circumstances for which he can claim no credit. The notion of desert seems not to apply to these cases' (104).

This is not to deny the role of individual entitlements altogether. Certain kinds of entitlements are consistent with the difference principle, but here it is necessary to distinguish between desert and legitimate expectations. Since it is in the general interest that I cultivate and exercise (some of) the talents and assets in my charge, rather than have them lie dormant, society is typically arranged to provide resources for their cultivation and incentives for their exercise. To be sure, I am entitled to my share of these benefits when I have qualified for them under the terms specified. What is important to stress, however, is that *these* claims honor the legitimate expectations created by institutions designed to elicit my efforts, not a primordial right or claim of desert in virtue of qualities I possess.

It is perfectly true that given a just system of co-operation as a scheme of

public rules and the expectations set up by it, those who, with the prospect of improving their condition, have done what the system announces that it will reward are entitled to their advantages. In this sense the more fortunate have a claim to their better condition; their claims are legitimate expectations established by social institutions, and the community is obligated to meet them. But this sense of desert presupposes the existence of the co-operative scheme; it is irrelevant to the question whether in the first place the scheme is to be designed in accordance with the difference principle or some other criterion (103).

Although I am *entitled* to the benefits answering my legitimate expectations, I do not *deserve* them, for two reasons: first, given the assumption of common assets, I do not really *possess* the attributes that give rise to the benefits, or if I do possess them, it is only in the weak, accidental sense rather than the strong, constitutive sense, and this sense of possession is inadequate to establish desert in the strong, pre-institutional sense. And second, while I am entitled to my fair share under the rules, I am not entitled that *these* rules, rewarding *these* attributes, be in force rather than some others. For these reasons, the well-endowed 'cannot say that he deserves and therefore has a right to a scheme of co-operation in which he is permitted to acquire benefits in ways that do not contribute to the welfare of others. There is no basis for his making this claim' (104).

MERITOCRACY VERSUS THE DIFFERENCE PRINCIPLE

Before moving on to consider Nozick's answer to Rawls and his defense of natural liberty, it may be helpful to clarify further some of the contrasts between the difference principle (as in Rawls' 'democratic equality') and the meritocratic conception (as in 'liberal equality'). Perhaps the most striking difference is in the role of individual merit or desert, which is central in the meritocratic conception and absent, or at least seriously weakened, in justice as fairness.¹ In a 'fair meritocracy', that is, one in which discrimination and class biases are overcome, those who achieve favored positions have earned their status and so deserve the rewards that attach to it. Unequal distributive shares are allocated in recognition of superior

¹ I leave aside those versions of meritocracy that would allocate distributive shares for the sake of creating incentives and attracting the relevant talents alone, without reference to the moral worthiness of the recipients.

achievement and not merely in satisfaction of legitimate expectations. As one defender of the meritocratic ethic explains,

A meritocracy is made up of those who have earned their authority.... Meritocracy, in the context of my usage, is an emphasis on individual achievement and earned status as confirmed by one's peers.... While all men are entitled to respect, they are not all entitled to praise. The meritocracy, in the best meaning of that word, is made up of those worthy of praise (Bell 1973: 453-4).

A second, related difference concerns the distinction between genetic and cultural advantages. For Rawls, this distinction is virtually irrelevant to questions of justice. On meritocratic assumptions, however, the distinction is crucial; hence the intense debate among those committed to meritocratic ideals over the role of genetic and cultural factors in determining intelligence and life prospects generally. Where the justice of distributive arrangements is seen to depend on the 'fair opportunity' of all to compete equally for (ultimately) unequal rewards, the distinction of genetic from social and cultural obstacles becomes central to any assessment of the scheme. The more closely success can be traced to hereditary factors, the fewer inequalities social institutions can be presumed capable of (or responsible for) ameliorating, and the less scope seems available for the kind of individual effort on which desert is said to depend.

In the nature of meritocracy, as it has been traditionally conceived, what is central to the assessment of a person is the assumed relation of achievement to intelligence, and of intelligence to its measurement on the Intelligence Quotient scale. The first question, therefore, is what determines intelligence.

All this makes the question of the relation of intelligence to genetic inheritance very touchy. Is intelligence largely inherited? Can one raise intelligence by nurture? How does one separate native ability and drive from improvements in skill acquired through education? (Bell 1973: 411)

For justice as fairness, the debate over the determinants of intelligence and the extensive social scientific literature it has produced are more or less beside the point. Once one rejects the notions of individual desert and 'fair opportunity' as the primary bases of distributive shares, the distinction between genetic and cultural obstacles to success loses much of its moral interest. Once we agree to regard the

distribution of talents as a common asset, it matters little how some came to reside in you and others in me.

Defenders of the meritocratic conceptions are not always explicit on the *grounds* for their distinction between natural and social advantages, but one can imagine at least two possible arguments, one moral and the other practical. The first would argue that genetic endowments are inviolable in a way that social or cultural characteristics are not, that a person's natural endowments are somehow more essentially *his*, more deeply constitutive of his identity than his socially-conditioned attributes. Innate differences, however arbitrary, are not dispensable in the same way; they, and not the results of social and cultural conditioning, mark out the traits in the absence of which I would not be the particular person I am. On this view, whether I deserve the intelligence with which I am born, for example, is not the point; what matters is that my native intelligence is a fact irreducibly about *me*, and the integrity of my person requires that it not be tampered with, no matter how worthy the wider social purpose.

But this argument seems flawed once it appears that those qualities most plausibly regarded as essential to a person's identity – one's character, values, core convictions, and deepest loyalties, for example – are often heavily influenced by social and cultural factors, while many natural features – such as hair color and other trivial physical characteristics – are more readily dispensable. Even if some distinction between essential and merely accidental characteristics of the person is valid, there seems no obvious reason why it must correspond to the distinction between natural and social assets. Of course, Rawls' objection to the argument would be stronger still, for his theory of the person implies that no characteristic, whether social or natural, can be essential in this way. Even those attributes, such as a person's character and values, that intuitively seem closest to defining an essential self, are relegated to contingent status. As a person's character 'depends in large part upon fortunate family and social circumstances for which he can claim no credit' (104), so our values are accidental as well. 'That we have one conception of the good rather than another is not relevant from a moral standpoint. In acquiring it we are influenced by the same sort of contingencies that lead us to rule out a knowledge of our sex and class' (Rawls 1975: 537).

The practical argument would distinguish natural and social inequalities on the grounds that the first are intractable in a way that the second are not, and that society can therefore be held responsible for remedying social inequalities but not natural ones; the more inequality turns out to be genetic rather than culturally-induced, the less society can 'do about it'. Given a fair system, some will advance more successfully than others, and there comes a point when even the most enlightened society can do nothing to alter this fact. At some point, even the most dogged reformers must acknowledge that life is unfair in a way that no social institutions can hope to set right. People are different, and these differences, sooner or later, will inevitably come to the fore, even – perhaps most assuredly – in a society where fair opportunity prevails. 'What is important is that the society, to the fullest possible extent, be a genuinely open one' (Bell 1973: 454).

To this Rawls would likely respond that society's role is naturally limited in this way only if one assumes that its only role in promoting justice consists in its efforts to even out the disadvantages of the least fortunate so that they may compete more fairly. But this assumption overlooks the equally significant social choice implicit in the aims institutions pursue and the attributes they reward in the process. Even if the vast majority of differences between persons turned out to be genetic rather than cultural, it would still remain for a society to determine *which* of these differences, if any, should properly be made the basis for differential distributive shares. To be sure, if the aim of social institutions is taken to be fixed, as maximizing the overall social product, for example, then the defenders of a 'fair meritocracy' are right, that the only remaining question of justice is how well people are equipped by the society to contribute to that purpose and to reap the benefits of their contribution. But the question must to reap the benefits of their contribution. But the question must immediately arise as to why *that* aim should be primary, even if it can prevail only at the admitted expense of social injustice. In short, a society not only equips or fails to equip its members with the endowments relevant to its collective purpose, but also defines through its institutional arrangements the nature of that purpose and, derivatively, the attributes to be prized and installed as the basis of distributive shares. In a passage of some eloquence, Rawls writes:

In view of these remarks we may reject the contention that the ordering of

institutions is always defective because the distribution of natural talents and the contingencies of social circumstances are unjust, and this injustice must inevitably carry over to human arrangements. Occasionally this reflection is offered as an excuse for ignoring injustice, as if the refusal to acquiesce in injustice is on a par with being unable to accept death. The natural distribution is neither just nor unjust; nor is it unjust that men are born into society at some particular position. These are simply natural facts. What is just and unjust is the way that institutions deal with these facts (102).

These considerations lead naturally to a third contrast between the meritocratic conception and the democratic one, which concerns the relation between the value of various assets and attributes on the one hand and the institutions that prize and reward them on the other. On the meritocratic conception, social institutions are bound to reward certain attributes rather than others. The qualities that a set of institutions calls forth have a worth antecedent to their institutional valuation, and so provide an independent test of the justice of the institutions themselves. Institutional arrangements that put a premium on noble qualities rather than base ones are, in virtue of *that*, the worthier, quite apart from other considerations relevant to justice, such as the purposes they advance.

On Rawls' view, institutions are not constrained in this way, for the virtues that would constrain them must themselves await institutional definition. 'The concept of moral worth does not provide a first principle of distributive justice', because it cannot be introduced until after the principles of justice are already on hand (312). Since no virtue has antecedent, or pre-institutional moral status, the design of institutions is open with respect to the qualities it may prize. As a result, the intrinsic worth of the attributes a society elicits and rewards cannot provide a measure for assessing its justice, for their worth only appears in the light of institutional arrangements to begin with. Rawls' rejection of pre-institutional notions of virtue reflects the priority of the right over the good and the refusal to choose in advance between competing conceptions of the good. A hunting society that rewards fleet-footedness over loquaciousness (as prized in a litigious society, for example) is not less just or worthy in virtue of *that*, for there are no antecedent grounds on which swiftness of foot could be shown as more or less virtuous than swiftness of tongue. The priority of just institutions with respect to

virtue and moral worth provides a second reason why I cannot be said to deserve the benefits flowing from my natural attributes. In order for me to deserve the benefits associated with 'my' superior intelligence, say, it is necessary both that I possess my intelligence (in some non-arbitrary sense of possession), *and* that I have a right (in a strong, pre-institutional sense of right) that society value intelligence rather than something else. But on Rawls' account, neither condition holds. The argument from arbitrariness to common assets undermines the first, and the precedence of institutions over moral worth denies the second.

DEFENDING COMMON ASSETS

In developing our critique of Rawls' theory of the person and the difference principle, each in the light of the other, we shall take as our points of departure two strands of Nozick's critique of justice as fairness. The first attacks the difference principle and in particular the notion of possession on which it relies, and the second defends a version of natural liberty by fixing on the notions of desert and entitlement. Playing Nozick off against Rawls in this way will enable us to assess Rawls' conception and also to clarify some of the similarities and differences in their respective distributive theories.

Nozick's central objection to the argument producing the difference principle fixes on Rawls' view that the distribution of natural talents is best regarded as a 'common' or 'collective' possession to be shared across society as a whole. As Rawls writes,

The difference principle represents, in effect, an agreement to regard the distribution of natural talents as a common asset and to share in the benefits of this distribution whatever it turns out to be (101).

The two principles are equivalent, as I have remarked, to an undertaking to regard the distribution of natural abilities as a collective asset so that the more fortunate are to benefit only in ways that help those who have lost out (179).

Rawls believes the notion of common assets as embodied in the difference principle expresses the ideal of mutual respect deontological liberalism seeks to affirm.

By arranging inequalities for reciprocal advantage and by abstaining from the exploitation of the contingencies of natural and social circumstance

within a framework of equal liberty, persons express their respect for one another in the very constitution of their society Another way of putting this is to say that the principles of justice manifest in the basic structure of society men's desire to treat one another not as means only but as ends in themselves (179).

Nozick argues, to the contrary, that to regard people's natural assets as common property is precisely to contradict all that deontological liberalism affirms in emphasizing the inviolability of the individual and the distinction between persons.

People will differ in how they view regarding natural talents as a common asset. Some will complain, echoing Rawls against utilitarianism, that this 'does not take seriously the distinction between persons'; and they will wonder whether any reconstruction of Kant that treats people's abilities and talents as resources for others can be adequate. 'The two principles of justice . . . rule out even the tendency to regard men as means to one another's welfare.' Only if one presses *very* hard on the distinction between men and their talents, assets, abilities and special traits (1974: 228).

Here Nozick goes to the heart of Rawls' theory of the subject. For as we have seen, Rawls presses very hard indeed on the distinction between the self and its various possessions. The severity of this distinction, problematic though it may be, is carefully fashioned to the requirements of the deontological project as a whole; it allows for the priority of the self over its ends, which supports in turn the priority of the right and the primacy of justice. A further feature of this conception is that it allows for the following defense against Nozick's objection to the difference principle.

Regarding the distribution of natural talents as a common asset does not violate the difference between persons nor regard some as means to others' welfare, because not *persons* but only 'their' *attributes* are being used as means to others' well-being. To say that *I* am somehow violated or abused when 'my' intelligence or even effort is used for the common benefit is to confuse the self with its contingently-given and wholly inessential attributes (inessential, that is, to me being the particular self I am). Only on a theory of the person that held these endowments to be essential *constituents* rather than alienable *attributes* of the self could the sharing of assets be viewed as using *me* as a means to others' ends. But on Rawls' account all endowments are contingent and in principle detachable from the self, whose

priority is assured by its ability constantly to recede before the swirl of circumstance. This is the feature that preserves its identity, by assuring its invulnerability to transformation by experience.

While this defense evades the inconsistency, it quickly invites a related objection of incoherence, for if Rawls must invoke the distinction between the self and its possessions in this thoroughgoing way, the question immediately arises whether, in avoiding a radically situated subject, Rawls does not lapse into the opposite extreme of a radically disembodied one. As Nozick presses the objection, 'Whether any coherent conception of a person remains when the distinction is so pressed is an open question. Why we, thick with particular traits, should be cheered that (only) the thus purified men within us are not regarded as means is also unclear' (1974: 228).

Nozick thus anticipates Rawls' defense and shows it to be too ingenious to redeem the theory. The notion that only *my* assets are being used as a means, not *me*, threatens to undermine the plausibility, even the coherence, of the very distinction it invokes. It suggests that on the difference principle, we can take seriously the distinction between persons only by taking metaphysically the distinction between a person and his attributes. But this has the consequence of leaving us with a subject *so* shorn of empirically-identifiable characteristics (so 'purified', in Nozick's word), as to resemble after all the Kantian transcendent or disembodied subject Rawls set out to avoid. It would seem that Rawls escapes the charge of inconsistency only at the price of incoherence, and that Nozick's objection to the difference principle therefore succeeds.

But Rawls has an alternate defense available, this one unanticipated by Nozick. Although it rescues the difference principle from reliance on an apparently disembodied conception of the subject, it comes at some expense to other aspects of Rawls' doctrine, and so would likely be resisted by Rawls himself. Still, I shall try to show that only by somehow assuming it can the difference principle be sustained. On this second defense, Rawls might deny that the difference principle uses me as a means to others' ends, not by claiming that my *assets* rather than *person* are being used, but instead by questioning the sense in which those who share in 'my' assets are properly described as 'others'. Where the first defense presses the distinction between the self and its attributes, the second qualifies the distinction between the self and the other by allowing that, in

certain moral circumstances, the relevant description of the self may embrace more than a single empirically-individuated human being. The second defense ties the notion of common assets to the possibility of a common subject of possession. It appeals, in short, to an intersubjective conception of the self.

That the difference principle commits Rawls to an intersubjective conception he otherwise rejects seems the only way out of the difficulties Nozick raises. It also serves to highlight an unargued assumption in Rawls' theory of the subject. As we have seen, Rawls conceives the self as a subject of possession, bounded in advance, and given prior to its ends, and he assumes furthermore that the bounds of the subject unproblematically correspond to the bodily bounds between individual human beings. But this claim is never defended by Rawls, only assumed, its contestability hidden perhaps by its affinity with our unreflective common-sense view of the matter.² Rawls' emphasis on plurality as an essential feature of human society may be thought to lend some support to the assumption, but this can establish only that *some* principle of plurality or differentiation is essential to an account of the human subject, not necessarily a physical bodily one; nor can it show that the number of this plurality must correspond in all cases to the number of individual human beings in the world.

This in any case is the assumption that must give way if Nozick's objection to common assets is to be overcome. If the difference principle is to avoid using some as means to others' ends, it can only be possible under circumstances where the subject of possession is a 'we' rather than an 'I', which circumstances imply in turn the existence of a community in the constitutive sense.

The conclusion that Rawls' theory implicitly relies on an intersubjective conception he officially rejects finds further support in the discussions of desert and justification to follow, and is hinted at by various traces of intersubjective language to be found throughout the text. Such language first appears in the discussion of the difference principle when, as we have seen, the distribution of natural talents is described alternately as a 'common' or 'collective' or 'social' asset to be used for the 'common advantage' (101, 179, 107).³ In justice as Still, Rawls' critique of utilitarianism makes especially perplexing his failure to defend his principle of individuation more completely, for it is precisely on this issue that he claims utilitarianism errs, by failing to recognize, or at any rate failing to take seriously, the distinction between persons.

fairness men agree to share one another's fate' (102). They resolve, that is, not to take seriously the differences between persons as the basis for differential life prospects, on the grounds that these differences arise from factors arbitrary from a moral point of view.

In his discussion of the idea of social union, Rawls carries his intersubjective language from common assets to common ends and purposes, and in rhetoric that comes perilously close to the teleological, speaks of human beings realizing their common nature as well. In his account of social union, Rawls shifts from distributive issues to a concern for self-realization, and seeks to show that justice as fairness can provide an interpretation of human sociability that is neither trivial nor purely instrumental. 'Thus human beings have in fact shared final ends and they value their common institutions and activities as goods in themselves' (522). The characteristics of social union are 'shared final ends and common activities valued for themselves' (525). Following Humboldt, a nineteenth century liberal in the German idealist tradition, Rawls writes that 'it is through social union founded upon the needs and potentialities of its members that each person can participate in the total sum of the realized natural assets of the others . . . *Only in the social union is the individual complete*' [emphasis added] (523, 525n). Social unions come in a variety of shapes and sizes; 'they range from families and friendships to much larger associations. Nor are there limits of time and space, for those widely separated by history and circumstance can nevertheless co-operate in *realizing their common nature*' [emphasis added] (527).

By their intersubjective dimensions, the difference principle and the idea of social union counter individualistic assumptions in two different ways, the difference principle by nullifying the arbitrariness that arises when natural assets are seen as individual possessions, the idea of social union by overcoming the partiality of persons that appears when individuals are thought to be complete in themselves. In a social union, 'the members of a community participate in one another's nature . . . [and] *the self is realized in the activities of many selves*' [emphasis added] (565).

It is a feature of human sociability that we are by ourselves but parts of what we might be. We must look to others to attain the excellences that we must leave aside, or lack altogether. The collective activity of society, the many associations and the public life of the largest community that regulates

them, sustains our efforts and elicits our contribution. Yet the good attained from the common culture far exceeds our work in the sense that we *cease to be mere fragments*: that *part of ourselves* that we directly realize is joined to a wider and just arrangement the aims of which we affirm [emphasis added] (529).

THE BASIS OF DESERT

The notion of possession leads naturally to claims of desert and entitlement. The argument over what people possess, and on what terms, has a direct bearing on the question of what people deserve or are entitled to as a matter of justice. It is to the issues of desert and entitlement that we now turn, to consider the second strand of Nozick's critique of justice as fairness. Rawls rejects the principles of natural liberty and liberal equality on the grounds that they reward assets and attributes which, being arbitrary from a moral point of view, people cannot properly be said to deserve, and adopts the difference principle on the grounds that it nullifies this arbitrariness. Nozick attacks this line of reasoning by arguing first that arbitrariness does not undermine desert, and second that, even if it did, a version of natural liberty and not the difference principle would emerge as the preferred result.

Stated in terms of possession, Rawls' objection to natural liberty and liberal equality is that under these principles, persons are allowed unfairly to benefit (or suffer) from natural and social endowments that do not properly *belong* to them, at least not in the strong, constitutive sense of belonging. To be sure, the various natural assets with which I am born may be said to 'belong' to me in the weak, contingent sense that they reside accidentally within me, but this sense of ownership or possession cannot establish that I have any special rights with respect to these assets or any privileged claim to the fruits of their exercise. In this attenuated sense of possession, I am not really the owner but merely the guardian or repository of the assorted assets and attributes located 'here'. By failing to acknowledge the arbitrariness of fortune, the principles of natural liberty and liberal equality go wrong in assuming that 'my' assets belong to me in the strong, constitutive sense, and so allowing distributive shares to depend on them.

Expressed in terms of desert, Rawls' objection to the principles of natural liberty and liberal equality is that they reward assets and

attributes that people cannot properly be said to deserve. Though some may think the fortunate deserve the things that lead to their greater advantage, 'this view is surely incorrect'.

It seems to be one of the fixed points of our considered judgments that no one deserves his place in the distribution of native endowments, any more than one deserves one's initial starting place in society. The assertion that a man deserves the superior character that enables him to make the effort to cultivate his abilities is equally problematic; for his character depends in large part upon fortunate family and social circumstances for which he can claim no credit. The notion of desert seems not to apply to these cases (104).

Because no one deserves his good luck in the genetic lottery, or his favored starting place in society, or for that matter the superior character that motivates him to cultivate his abilities conscientiously, no one can be said to deserve the benefits these assets produce. It is this deduction that Nozick disputes. 'It is not true,' he argues, 'that a person earns Y (a right to keep a painting he's made, praise for writing *A Theory of Justice*, and so on) only if he's earned (or otherwise *deserves*) whatever he used (including natural assets) in the process of earning Y. Some of the things he uses he just may *have*, not illegitimately. It needn't be that the foundations underlying desert are themselves deserved, *all the way down*' (1974: 225).

Now what are we to make of this claim? If I do not necessarily have to *deserve* everything I use in producing a thing in order to deserve the thing, what *does* my desert depend on? Nozick says that some of the things I use I 'just may *have*, not illegitimately' (and, presumably, possibly arbitrarily). Once again, the notion of possession enters the scene. To see whether my having a thing, not illegitimately, can enable me to deserve what it helps me produce, we must explore in greater detail the relation between possession and desert, and sort out once more the sense of possession being appealed to.

For this purpose, it may be helpful to consider a recent discussion of justice and personal desert by Joel Feinberg, who analyzes the bases of desert with an admirable clarity in terms suggestive for the arguments before us (1970). Feinberg begins with the observation that no one can deserve anything unless there is some basis for the desert. 'Desert without a basis is simply not desert'. But the question immediately arises what *kind* of basis is necessary. As Feinberg writes, 'Not any old basis will do'. Once again, the notion of posses-

sion provides the key. 'If a person is deserving of some sort of treatment, he must, necessarily, be so in virtue of some *possessed characteristic* or prior activity' [emphasis added] (1970: 48).

A characteristic of mine cannot be a basis for a desert of yours unless it somehow reveals or reflects some characteristic of yours. In general, the facts which constitute the basis of a subject's desert must be facts about that subject. If a student deserves a high grade in a course, for example, his desert must be in virtue of some fact about *him* — his earlier performances, say, or his present abilities It is necessary that a person's desert have a basis and that the basis consist in some fact about himself (1970: 58-9, 61).

Feinberg's analysis, tying a person's desert to some fact about the person, would appear to support Nozick's claim that 'the foundations underlying desert needn't themselves be deserved, *all the way down*'. In fact, the reliance of desert on some possessed characteristic of the person suggests a thesis even stronger than Nozick's: that the foundations underlying desert *cannot* themselves be deserved, *all the way down*, any more than the foundations underlying possession can themselves be possessed, *all the way down*. We have already seen how the notion of possession requires that somewhere, 'down there', there must be a subject of possession that is not *itself* possessed (for this would deny its agency), a subject 'doing the possessing', so to speak. The analogy for desert must be a *basis* of desert ultimately prior to desert. For consider: if desert presupposes some possessed characteristic, and if possessed characteristics presuppose some subject of possession which is not itself possessed, then desert must presuppose some subject of possession which is not itself possessed, and therefore some basis of desert which is not itself deserved. Just as there must be some subject of possession prior to possession, so there must be some basis of desert prior to desert. This is why the question whether someone deserves (to have) his sterling character, for example, is notoriously difficult (for it is unclear who or what is left to judge once his character has been removed), and why, beyond a certain point, asking just wholesale whether someone deserves to be the (kind of) person he is becomes incoherent altogether. Somewhere, 'down there', there must be a basis of desert that is not itself deserved. The foundations underlying desert cannot themselves be deserved, all the way down.

This result would seem amply to confirm Nozick's claim against Rawls that I do not necessarily have to *deserve* everything I use in producing a thing in order to deserve the thing, that some of what I use I 'just may *have*, not illegitimately'. And if this claim can be established, then it would appear that Rawls' argument from arbitrariness fails to undermine desert after all. To say, as Rawls does, that I do not deserve the superior character that led me to realize my abilities is no longer enough. To deny my desert, he must show that I do not *have* the requisite character, or alternatively, that I *have* it, but not in the requisite sense.

But this is precisely the argument Rawls' theory of the person allows him to make. For given his sharp distinction between the self, taken as the pure subject of possession, and the aims and attributes it possesses, the self is left bare of any substantive feature or characteristic that could qualify as a desert base. Given the distancing aspect of possession, the self *itself* is dispossessed. On Rawls' theory of the person, the self, strictly speaking, *has nothing*, nothing at least in the strong, constitutive sense necessary to desert. In a move similar to the one invoked to show that the difference principle does not use a *person* as a means, only a person's *attributes*, Rawls can accept that some undeserved desert base is necessary to desert, only to claim that, on an adequate understanding of the person, this condition could never in principle be met! On Rawls' conception, the characteristics I possess do not *attach* to the self but are only *related* to the self, standing always at a certain distance. This is what makes them attributes rather than constituents of my person; they are *mine* rather than *me*, things I *have* rather than *am*.

We can see in this light how Rawls' argument from arbitrariness undermines desert not directly, by claiming I cannot *deserve* what is arbitrarily given, but indirectly, by showing I cannot *possess* what is arbitrarily given, that is, that 'I', *qua* subject of possession, cannot possess it in the undistanced, constitutive sense necessary to provide a desert base. An arbitrarily-given asset cannot be an essential constituent but only an accidental attribute of my person, for otherwise my identity would hang on a mere contingency, its continuity constantly vulnerable to transformation by experience, my status as a sovereign agent dependent on the conditions of my existence rather than epistemologically guaranteed. On Rawls' conception, no one can properly be said to deserve anything because no one can properly

ly be said to possess anything, at least not in the strong, constitutive sense of possession necessary to the notion of desert.

A theory of justice without desert would seem a dramatic departure from traditional conceptions, but Rawls is at pains to show that it is not. In his opening pages, Rawls acknowledges that his approach 'may not seem to tally with tradition', but seeks to reassure that in fact it does.

The more specific sense that Aristotle gives to justice, and from which the most familiar formulations derive, is that of refraining from *pleonexia*, that is, from gaining some advantage for oneself by seizing *what belongs to another*, his property, his reward, his office, and the like, or by denying a person that which is due to him *Aristotle's definition clearly presupposes, however, an account of what property belongs to a person, and of what is due to him. Now such entitlements are, I believe, very often derived from social institutions and the legitimate expectations to which they give rise.* There is no reason to think that Aristotle would disagree with this, and certainly he has a conception of social justice to account for these claims There is no conflict with the traditional notion [emphasis added] (10-11).

In comparing justice as fairness with traditional conceptions, Rawls confirms its novelty rather than denies it. What he presents as an incidental qualification to justice as classically conceived turns out on inspection to signal a striking departure. As Rawls suggests, traditional notions freely refer to 'what properly belongs to a person', institutions, presumably, aside; they presuppose thickly-constituted persons with a fixity of character, certain features of which are taken to be essential, 'all the way down'. On Rawls' conception, however, none of these concepts is available. In so far as a theory of justice 'presupposes an account of what properly belongs to a person' (in the strong sense of 'belongs'), Rawls effectively acknowledges that he has none. Nor, he seems to imply, given the precedence of plurality, the priority of right, and the theory of the person they require, is it reasonable to think that such a theory of justice could be true. We are not essentially thick enough selves to bear rights and deserts antecedent to the institutions that define them. Given these constraints, the only alternative is to opt for a theory of justice based on entitlements to legitimate expectations, ruling out desert altogether. Rawls hedges this claim at first, saying only that 'such entitlements are, I believe, *very often* derived from social institutions and the legitimate expectations to which they give rise' [emphasis added] (10). But as

the full consequences of Rawls' view emerge, 'very often' becomes 'always', for it becomes clear that 'such entitlements' can arise in no other way. While Aristotle might not disagree that entitlements can arise in this way, it seems far from his view that they can arise in no other way. In denying that justice has to do with giving people what they deserve, justice as fairness departs decisively from the traditional notion after all.

Rawls' apparent view that no one can properly be said to deserve anything, and the connection of this view with the notion of the self as 'essentially unencumbered', emerges more fully in his discussion of legitimate expectations and moral desert. He begins by acknowledging that justice as fairness, in rejecting desert, runs counter to common sense.

There is a tendency for common sense to suppose that income and wealth, and the good things in life generally, should be distributed according to moral desert. Justice is happiness according to virtue. While it is recognized that this ideal can never be fully carried out, it is the appropriate conception of distributive justice, at least as a *prima facie* principle, and society should try to realize it as circumstances permit. Now justice as fairness rejects this conception. Such a principle would not be chosen in the original position. There seems to be no way of defining the requisite criterion in that situation (310-11).

There seems to be no way of defining the requisite criterion of a person's virtue or moral worth in the original position because no substantive theory of the person antecedent to social institutions exists. For moral desert to provide an independent criterion of justice, there must be some substantive theory of the person, or of the worth of persons, to get it going. But for Rawls, the worth of persons is subsequent to institutions, not independent of them. And so a person's moral claims must await their arrival.

This leads to the distinction between moral desert and legitimate expectations. Once a person does the various things established institutions encourage him to do, he acquires certain rights, but not before. He is entitled that institutions honor the claims they announce they will reward, but he is not entitled that they undertake to reward any particular kind of claim in the first place.

A just scheme, then, answers to what men are entitled to; it satisfies their legitimate expectations as founded upon social institutions. But what they

are entitled to is not proportional to nor dependent upon their intrinsic worth. The principles of justice that regulate the basic structure and specify the duties and obligations of individuals do not mention moral desert, and there is no tendency for distributive shares to correspond to it (311).

The principles of justice do not mention moral desert because, strictly speaking, no one can be said to deserve anything. Similarly, the reason people's entitlements are not proportional to nor dependent upon their intrinsic worth is that, on Rawls' view, *people have no intrinsic worth*, no worth that is intrinsic in the sense that it is theirs prior to or independent of or apart from what just institutions attribute to them.

The essential point is that the concept of moral worth does not provide a first principle of distributive justice. This is because it cannot be introduced until after the principles of justice and of natural duty and obligation have been acknowledged [T]he concept of moral worth is secondary to those of right and justice, and it plays no role in the substantive definition of distributive shares (312-13).

Rawls could agree with Feinberg that 'desert is a *moral* concept in the sense that it is logically prior to and independent of public institutions and their rules', but would deny that there is any 'antecedent standard for its definition' (313), and so disagree with Feinberg that 'one of the aims of [a system of public bestowals] is to give people what they deserve' (1970: 86). For Rawls, the principles of justice aim neither at rewarding virtue nor at giving people what they deserve, but instead at calling forth the resources and talents necessary to serve the common interest.

None of the precepts of justice aims at rewarding virtue. The premiums earned by scarce natural talents, for example, are to cover the costs of training and to encourage the efforts of learning, as well as to direct ability to where it best furthers the common interest. The distributive shares that result do not correlate with moral worth (311).

To illustrate the priority of just institutions with respect to virtue and moral worth, Rawls suggests an analogy to the relation between the rules of property and the law of robbery and theft.

These offenses and the demerits they entail presuppose the institution of property which is established for prior and independent social ends. For a society to organize itself with the aim of rewarding moral desert as a first

principle would be like having the institution of property in order to punish thieves. The criterion to each according to his virtue would not, then, be chosen in the original position (313).

The analogy is intriguing, but one wonders whether it works entirely to Rawls' advantage. While it is apparent that the institution of property has a *certain* priority with respect to its correlative offenses, it is less clear why the dependence must run only in one direction, especially given Rawls' own commitment to the method of the reflective equilibrium. For example, is our belief in the validity of the institution of property in no way enhanced by a conviction that robbery and theft are wrong? Would our confidence in the institution of property in no way be diminished if it turned out that those it defined as robbers and thieves were invariably good and virtuous men? And what of more extreme cases? While the norms and rules of protecting human life can no doubt be defended on a variety of grounds, such as keeping people alive, avoiding suffering, and so on, is it logically mistaken to think that one justification of prohibitions against murder could be to punish murderers?

Rawls' position here appears especially perplexing in the light of a contrast he draws between distributive justice and retributive justice, suggesting that in the second case, some notion of moral desert may be appropriate after all. The view that distributive shares should match moral worth to the extent possible, writes Rawls, 'may arise from thinking of distributive justice as somehow the opposite of retributive justice'. But the analogy is mistaken. In a reasonably well-ordered society, 'Those who are punished for violating just laws have normally done something wrong. This is because the purpose of the criminal law is to uphold basic natural duties . . . and punishments are to serve this end'.

They are not simply a scheme of taxes and burdens designed to put a price on certain forms of conduct and in this way to guide men's conduct for mutual advantage. It would be far better if the acts prescribed by penal statutes were never done. *Thus a propensity to commit such acts is a mark of bad character*, and in a just society legal punishments will only fall upon those who display these faults.

It is clear that the distribution of economic and social advantages is entirely different. These arrangements are not the converse, so to speak, of the criminal law so that just as the one punishes certain offenses, the other rewards moral worth. The function of unequal distributive shares is to

cover the costs of training and education, to attract individuals to places and associations where they are most needed from a social point of view, and so on To think of distributive and retributive justice as *converses of one another is completely misleading and suggests a moral basis of distributive shares where none exists* [emphasis added] (314-15).

Unlike the benefits that flow from distributive arrangements, the punishments and prohibitions associated with the criminal law are not simply a non-moral system of incentives and deterrents designed to encourage some forms of behavior and discourage others. For Rawls, the pre-institutional moral notions excluded in distributive justice somehow find meaning for retributive purposes, and there is a tendency for punishment to correspond to them.

The immediate puzzle is how this account can possibly fit with the analogy of property and theft. If retributive justice differs from distributive justice precisely in virtue of its prior moral basis, it is difficult to see how the example of property and theft could demonstrate the priority of social institutions with respect to virtue and moral worth, if this priority holds for distributive justice alone. This relatively minor confusion aside, the more basic question is how Rawls can admit desert in retributive justice without contradicting the theory of the self and related assumptions that ruled it out for purposes of distributive justice. If such notions as pre-institutional moral claims and intrinsic moral worth are excluded from a theory of distributive justice in virtue of an essentially unencumbered self too slender to support them, it is difficult to see how retributive justice could differ in any relevant way.³

Do not the same arguments from arbitrariness exclude desert as a basis for punishment as for distributive shares? Is the propensity to commit crimes, any less than the propensity to do good, the result of factors arbitrary from a moral point of view? And if not, why would the parties to the original position not agree to share one another's fate for the purpose of criminal liability as well as distributive arrangements? Since under the veil of ignorance, none can know whether he shall have the misfortune to be born into the unfavorable social and family

³ In a footnote, Rawls (315) cites Feinberg in apparent support of this claim, but Feinberg allows a role for desert in both distributive and retributive justice. Feinberg's point is that retributive justice involves what he calls polar desert (where one either deserves good or deserves ill), whereas distributive justice involves nonpolar desert (where, as with a prize, some deserve and others do not). But both cases involve desert in the moral, pre-institutional sense (Feinberg 1970: 62).

circumstances that lead to a life of crime, why would the parties not adopt a kind of difference principle for punishments as well as distributive shares, and agree, in effect, to regard the distribution of natural and social liabilities as a common burden?

Rawls holds that 'those who are punished for violating just laws have normally done something wrong', and so deserve their punishment (314). But suppose, by an act of vandalism, I deprive the community of a certain measure of well-being, say by throwing a brick through a window. Is there any reason why I deserve to bear the full costs of my destructiveness any more than the person who produced the window *deserves* to enjoy the full benefits of his productiveness? Rawls may reply that my 'propensity to commit such acts is a mark of bad character'. But if the worker's industriousness in making the window is not a mark of good character (in the moral, pre-institutional sense), why is my maliciousness in breaking the window a mark of bad character (in the moral, pre-institutional sense)? To be sure (following Rawls, p. 103), given a just system of criminal law, those who have done what the system announces it will punish are properly dealt with accordingly and in this sense are 'deserving' of their penalty. But this sense of desert presupposes the existence of the [retributive] scheme; it is irrelevant to the question whether in the first place the scheme is to be designed in accordance with the difference principle or some other criterion' (103).

Some may think that the criminal deserves his punishment in the strong moral sense because he deserves the low character his criminality reflects. Perhaps this is what Rawls has in mind when he writes that 'propensity to commit such acts is a mark of bad character', and punishments properly fall on those who display these faults. Because the transgressor is less worthy in this sense, he deserves the misfortune that befalls him. But again (following Rawls, p. 104), this view is surely incorrect. It seems to be one of the fixed points of our considered judgments that no one deserves his place in the distribution of native endowments or liabilities, any more than one deserves one's initial starting place in society. The assertion that a man deserves the inferior character that prevents him from overcoming his liabilities is equally problematic; for his character depends in large part upon unfortunate family and social circumstances for which he cannot be blamed. The notion of desert seems not to apply to these cases. None of which is to say that, generally

speaking, a non-moral theory of distributive justice is incompatible with a moral, or desert-based theory of punishment, only that given Rawls' reasons for rejecting desert-based distributive arrangements, he seems clearly committed to rejecting desert-based retributive ones as well.

The apparent inconsistency between Rawls' retributive and distributive theories need not do serious damage to the theory as a whole. Given the method of reflective equilibrium, 'justification is a matter of the mutual support of many considerations, of everything fitting together into one coherent view' (21). From the standpoint of the overall theory, little hangs on Rawls' retributive theory, apart from the measure of plausibility it lends justice as fairness for those committed to a strong, desert-based notion of punishment. If Rawls' distinction succeeds, they need not choose between their retributive intuitions and the difference principle; if it does not, one or the other of those convictions must give way. If, on reflection, a non-moral theory of punishment appears unacceptable, even in the light of the arbitrariness of criminal characteristics and dispositions, then the difference principle – rejecting as it does the notion of desert – would be called into serious question. If, on the other hand, our intuition that criminals deserve punishment proves no more indispensable than our intuition that virtue deserves reward (an intuition of common sense Rawls explicitly rejects), then we may adjust our intuitions in a direction that affirms the difference principle rather than opposes it. Desert would be rejected as the basis for both distributive and retributive arrangements, and so the inconsistency resolved.

But such a resolution returns us to the larger difficulties of a theory of justice without desert and a notion of the self as essentially disposessed, or barren of constituent traits. Nozick argues against Rawls that the foundations underlying desert need not themselves be deserved, all the way down. But as we have seen, Rawls' denial of desert does not depend on the thesis Nozick refutes, but instead on the notion of the self as a pure, unadulterated, 'essentially unencumbered' subject of possession. Rawls is not committed to the view that a person can only deserve a thing he produces if he deserves everything he used in producing it, but rather to the view that no one possesses anything in the strong, constitutive sense necessary to a desert base. No one can be said to deserve anything (in the strong,

pre-institutional sense), because no one can be said to possess anything (in the strong, constitutive sense). This is the philosophical force of the argument from arbitrariness.

That the argument from arbitrariness works in this way can be seen by viewing the moves from natural liberty to fair opportunity to the democratic conception, as traced by Rawls, as stages in the dispossession of the person. With each transition, a substantive self, thick with particular traits, is progressively shorn of characteristics once taken to be essential to its identity; as more of its features are seen to be arbitrarily given, they are relegated from presumed constituents to mere attributes of the self. More becomes *mine*, and less remains *me*, to recall our earlier formulation, until ultimately the self is purged of empirical constituents altogether, and transformed into a condition of agency standing beyond the objects of its possession. The logic of Rawls' argument might be reconstructed as follows:

At the far end of the spectrum, even before natural liberty appears, are aristocratic and caste societies; in such societies, a person's life prospects are tied to a hierarchy into which he is born and from which his person is inseparable. Here, the self is most fully ascribed, merged almost indistinguishably with its condition, embedded in its situation. The system of natural liberty removes fixed status of birth as an assumed constituent of the person, and regards each as free, given his capacities and resources, to compete in the marketplace as best he can, and to reap his reward. By shifting the basis of expectations from status to contract, the system of natural liberty repairs the arbitrariness of hierarchical societies by taking the person more narrowly, so to speak, as distinct and separable from his surroundings. Still, some arbitrariness remains, most notably in the form of social and cultural contingencies. In the regime of natural liberty, a person's life prospects are governed by factors no more ascribable to the person (in the strong, constitutive sense) than his inherited status. Having relieved the person of his hierarchical baggage, the principle of natural liberty still conceives a thickly-constituted self, burdened by the accidents of social and cultural contingency. And so the move to fair opportunity, which strips the self of social and cultural accidents as well as inherited status. In a 'fair meritocracy', the effects of class status and cultural disadvantage are understood to reflect more on the society and less on the person. Those with

comparable talents and 'the same willingness to use them, should have the same prospects of success regardless of their initial place in the social system, that is, irrespective of the income class into which they are born' (73). In this way, the meritocratic conception extends the logic of natural liberty by ascribing less to the self and more to its situation.

But even the principle of fair opportunity, in rewarding individual effort, conceives the province of the self too expansively. For even 'the effort a person is willing to make is influenced by his natural abilities and skills and the alternatives open to him. The better endowed are more likely, other things equal, to strive conscientiously, and there seems to be no way to discount for their greater good fortune' (312). The self is still over-ascribed. Given its arbitrariness, even the character that determines a person's motivation cannot properly be regarded as an essential constituent of his identity. And so finally the move to the democratic conception, in which the self, shorn of all contingently-given attributes, assumes a kind of supra-empirical status, essentially unencumbered, bounded in advance and given prior to its ends, a pure subject of agency and possession, ultimately thin. Not only my character but even my values and deepest convictions are relegated to the contingent, as features of my condition rather than as constituents of my person. 'That we have one conception of the good rather than another is not relevant from a moral standpoint. In acquiring it we are influenced by the same sort of contingencies that lead us to rule out a knowledge of our sex and class' (Rawls 1975: 537). Only in this way is it possible to install the self as invulnerable, to assure its sovereignty once and for all in a world threatening always to engulf it. Only if the fate of the self is thus detached from the fate of its attributes and aims, subject as they are to the vagaries of circumstance, can its priority be preserved and its agency guaranteed.

This is the vision of the person that Nozick and Bell, as defenders of natural liberty and meritocracy, respectively, emphatically reject, even if they do not spell out in any detail the conception of the self they rely on instead. Both object that the argument from arbitrariness, consistently applied, leads ineluctably to the dissolution of the person, and the abnegation of individual responsibility and moral choice. 'This line of argument can succeed in blocking the introduction of a person's autonomous choices and activities (and their

results) only by attributing *everything* noteworthy about the person completely to certain sorts of "external" factors', writes Nozick. Echoing his argument against the notion of common assets, Nozick questions whether, on Rawls' account, any coherent conception of the person remains, and if so, whether it is any longer the kind of person worth the moral fuss deontological liberalism makes on its behalf.

So denigrating a person's autonomy and prime responsibility for his actions is a risky line to take for a theory that otherwise wishes to buttress the dignity and self-respect of autonomous beings; especially for a theory that founds so much (including a theory of the good) upon a person's choices. One doubts that the unexalted picture of human beings Rawls' theory presupposes and rests upon can be made to fit together with the view of human dignity it is designed to lead to and embody (1974: 214).

Bell summarizes the objection in an epigram: 'The person has disappeared. Only attributes remain' (1973: 419). Where Rawls seeks to assure the autonomy of the self by disengaging it from the world, his critics say he ends by dissolving the self in order to preserve it.

To recapitulate our reconstructed version of the argument between Rawls and Nozick on the issue of desert: Nozick first argues that the arbitrariness of assets does not undermine desert, because desert may depend not only on things I deserve, but also on things I *have*, not illegitimately. Rawls' response is to invoke the distinction between the self and its possessions in the strongest version of that distinction, and so to claim that, strictly speaking, there is nothing that 'I', *qua* pure subject of possession, *have* – nothing that is attached, rather than related, to *me* – nothing at least in the strong, constitutive sense of possession necessary to a desert base. Nozick's rejoinder is that this defense cannot succeed for long, for it has the consequence of leaving us with a subject *so* shorn of empirically-identifiable characteristics as to resemble once more the Kantian transcendent or disembodied subject Rawls resolved to avoid. It makes the individual inviolable only by making him invisible, and calls into question the dignity and autonomy this liberalism seeks above all to secure.

INDIVIDUAL AND SOCIAL CLAIMS: WHO OWNS WHAT?

But Nozick has a further objection, independent of the first, that goes something like this: even if Rawls is right that arbitrariness undermines individual possession and hence individual desert, the difference principle is not the inevitable result; something like the principle of natural liberty – and here Nozick prefers to speak in terms of his own version, which he calls the ‘entitlement theory’ – could still be true. For even if no one deserved any of his natural assets, it might still be that people were entitled to them, and to what flows from them.

At best, this argument maintains, Rawls’ case for the difference principle is underdetermined by the argument from arbitrariness. To show that individuals, as individuals, do not deserve or possess ‘their’ assets is not necessarily to show that society as a whole *does* deserve or possess them. Simply because the attributes accidentally located in me are not *my* assets, why must it follow, as Rawls seems to think, that they are common assets, rather than nobody’s assets? If they cannot properly be said to belong to me, why assume automatically that they belong to the community? Is their location in the *community’s* province any less accidental, any less arbitrary from a moral point of view? And if not, why not regard them as free-floating assets, unattached in advance to *any* subject of possession, whether individual or social?

Here it is necessary to be more precise about the terms of relation between the person and the endowments he bears. Three descriptions seem to cover the possible cases; depending on the sense of possession intended, I may be described as the owner, the guardian, or the repository of the endowments I bear. In its strongest version, the notion of ownership implies that I have absolute, unqualified, exclusive rights with respect to my endowments, in its more moderate versions that I have certain privileged claims with respect to them, a bundle of rights, while not unlimited, at least more extensive with respect to my assets than any bundle of rights anyone else may have with respect to them. This is the sense of individual possession against which Rawls’ argument from arbitrariness is addressed and which, if successful, it undermines.

If Rawls is right, and I cannot properly be described as the owner of my assets and attributes, two alternate accounts remain. One is the notion of guardianship, which denies individual ownership in

favor of a more ultimate owner or subject of possession of which the individual person is the agent. To say that I am the guardian of the endowments I bear is to imply that they are owned by some other subject, on whose behalf, or in whose name, or by whose grace I cultivate and exercise them. This is a notion of possession reminiscent of the early Christian notion of property, in which man had what he had as the guardian of assets belonging truly to God, and it is a notion that fits with various communitarian notions of property as well.⁴

In the third account, I am neither the owner nor the guardian but rather the repository of the assets and attributes accidentally located in my person. The notion of the individual as repository of endowments does not presuppose some other subject of possession whose endowments they ultimately are, but does without possession altogether. In so far as I am the repository of natural assets, there need be no further question to whom these assets ultimately *belong*; their residing in me has no consequence for claims I, or others, may have with respect to them.

Now in terms of these distinctions, Nozick claims in effect that Rawls’ argument from arbitrariness, even if it succeeds in undermining individual possession and hence desert, warrants only the third description (that individuals are repositories of assets) and not the second. But if this is all the argument from arbitrariness establishes, then it does not lead to the difference principle. For the difference principle must presuppose the second description – that I am the guardian of assets to which the community as a whole has some prior title or claim. If all arbitrariness means is that I am the repository of assets which belong to no one in particular, then it cannot be assumed that the community owns them any more than I do. It

4 The Rev. Vernon Bartlet (1915: 97-8) writes that ‘the essential Christian attitude’ held the property rights of any individual to be ‘purely relative, not only as compared with God’s absolute rights as Producer and Owner both of all things and of all persons, but also as compared with the paramount human or derivative rights of Society as representing the common weal. Of this, the individual’s weal is only a dependent part, and should be limited by the rights of all others to the conditions of personal well-being....’ The resulting practical principles, viz. the stewardship of property on behalf both of God and Society, and the moral duty of fidelity in this relation as the condition of any correlative rights of private personal enjoyment, is too deeply embedded in Christ’s teaching, notably in the parables, to need demonstration.’ Bartlet quotes St. Paul, I *Corinthians* iv. 71: ‘For who maketh thee to excel? And what hast thou that thou hast not received? But if thou didst receive it, why dost thou glory as if thou hadst not received it?’

would be as though assets, and the benefits that flow from them, fell like manna from heaven (the image is Nozick's), unmarked by prior claims and unattached to any subject of possession whether individual or social.

How then should they be distributed? On what basis should competing claims to such spontaneously-generated bounty be assessed? From the standpoint of desert, there would seem to be no grounds on which to choose between letting the assets lie where they fall, and trying to distribute them in some other way. Unless distribution under such assumptions is to be regarded as a stand-off, a matter of moral indifference, people's entitlements must depend on considerations other than notions of possession or desert. On this, Rawls and Nozick seem prepared to agree. But what could these considerations be? On this, they part ways. Each thinks he can adduce considerations unrelated to desert to overcome the stand-off in favor of his own conception.

For Nozick, the absence of desert creates a presumption in favor of letting assets lie where they fall, at least once it is accepted that things do not come into the world like manna from heaven but come into being already held, attached to particular persons.

Since things come into being already held (or with agreements already made about how they are to be held), there is no need to search for some pattern for unheld holdings to fit The situation is not an appropriate one for wondering, 'After all, what is to become of these things; what are we to do with them.' In the non-manna-from-heaven world in which things have to be made or produced or transformed by people, there is no separate process of distribution for a theory of distribution to be a theory of (1974: 219).

Nozick goes on to argue that if a person has an asset to which no one else is entitled, then, although he may not *deserve* the asset, he is nonetheless *entitled* to it, and to whatever flows from it by a process that does not violate anyone else's entitlements. 'Whether or not people's natural assets are arbitrary from a moral point of view, they are entitled to them, and to what flows from them' (1974: 226).

For Rawls, on the other hand, the absence of individual desert creates a presumption in favor of regarding the distribution of talents as a common asset. The lack of desert or a pre-institutional concept of virtue means that institutions are unconstrained by antecedent

moral claims in their pursuit of the primary virtue of social justice. In this sense, the analogy of manna from heaven is apt. The array of assets dealt by fortune is neither just nor unjust. 'These are simply natural facts. What is just and unjust is the way that institutions deal with these facts' (102). There is no reason to let assets and the benefits that flow from them lie where they fall. This would be simply to incorporate and affirm the arbitrariness of nature. The discovery that virtue and entitlements await social institutions rather than constrain them is a reason to pursue justice all the more insistently, not a reason to freeze arbitrariness in place.

What, then, are we to make of these attempts to overcome the apparent moral stand-off created by the presumed absence of desert? Nozick is prepared to accept that people may not deserve their natural assets, but claims they are entitled to them nonetheless. 'If people have X, and their having X (whether or not they deserve to have it) does not violate anyone else's (Lockean) right or entitlement to X, and Y flows from (arises out of, and so on) X by a process that does not itself violate anyone's (Lockean) rights or entitlements, then the person is entitled to Y' (1974: 225). But he does not show why this is so, nor is he clear on what precisely the difference between desert and entitlement consists in. Rawls and Feinberg agree that desert is a *moral* concept in the sense that it is logically prior to and independent of public institutions and their rules' (Feinberg 1970: 87). Entitlements, by contrast, are claims that can only arise under the rules or qualifying conditions of institutions already established, what Rawls describes as legitimate expectations founded on social institutions. It is a consequence of this view that, for the purpose of designing or assessing social institutions, people's entitlements, being *derivative* from institutions, are without moral or critical force. Assessing the justice of an institution in the light of what people were entitled to would be like judging the validity of a rule in the light of claims arising under the rule; to recall our earlier discussion, it would be appealing to a standard of appraisal thoroughly implicated in the object of appraisal. It is for this reason that the concept of entitlement cannot provide a first principle of justice. As Rawls explains, it 'presupposes the existence of the co-operative scheme, and so is irrelevant to the question whether in the first place the scheme is to be designed in accordance with the difference principle or some other criterion' (103).

Nozick never comes to terms with this difficulty. In making his argument, he explicitly adopts the language of entitlement rather than desert, but does not acknowledge its lesser moral force. In Nozick's usage, the concept of entitlement does the same work as desert, but without its pre-institutional credentials ever being established. He begins with the premise that 'people are entitled to their natural assets', and proceeds to argue that they are entitled to the benefits that flow from them (1974: 225-6). But he never says why people are entitled to their assets in any sense of entitlement strong enough to get the argument going.

At one point, Nozick seems to claim that people are entitled to the things they 'just may have, not illegitimately', their natural assets presumably being among such things. But having something, not illegitimately, is not the same as being entitled to it; it is simply having it, in some unspecified sense of possession. While my intelligence, or physical strength, or good health may be among the things I have, not illegitimately, it does not follow that I am entitled to these things, for entitlements depend, as we have seen, on some scheme of co-operation already being in effect.

At other points, Nozick seems to argue instead that people are entitled to their natural assets and the benefits that flow from them in some sense of entitlement antecedent to social institutions, which is to say in some sense of entitlement equivalent to desert. ('It needn't be that the foundations underlying desert are themselves deserved, *all the way down.*') This would require that I possess the relevant assets in the strong sense of possession adequate to a desert base, and this in turn requires a theory of the person on which I possess some things, at least, as constituents and not merely as attributes of the self. But Nozick's theory of the person is not easy to discern. He complains that the 'purified' self implicit in Rawls' theory seems radically at odds with our more familiar notion of ourselves as beings 'thick with particular traits' (1974: 228), and further objects that so 'purified' a self, even if coherent, seems to threaten not only individual desert, but also such indispensable notions as individual autonomy, responsibility, and the very human dignity and self-respect deontological liberalism sets out to affirm (1974: 214). These, to be sure, are powerful objections. But they do little to articulate a substantive theory of the person that manages to leave the person 'thick with particular traits' while at the same time avoiding the

moral and epistemological difficulties Rawls identifies with the self laden with contingencies. It is one thing simply to assert what is in some sense undeniable, that we are 'thick with particular traits', and quite another to show how this can be true in a way not subject to the rival incoherences associated with a radically situated self, indefinitely conditioned by its surroundings and constantly subject to transformation by experience. For this reason, Nozick's proposed solution to the stand-off can be of critical interest alone.

What, then, of Rawls' attempt? If Nozick fails to show that the absence of individual desert leads to a presumption in favor of letting natural assets lie where they fall, with what success does Rawls show a presumption in favor of a general claim to such assets? According to Nozick, Rawls' view is that 'everyone has some entitlement or claim on the totality of natural assets (viewed as a pool), with no one having differential claims' (1974 228), and indeed the notion of common assets seems to imply some such view. The question is what kind of 'entitlement or claim' is involved, and how it might be established. Two possibilities suggest themselves; the general social claim on the totality of natural assets might be a claim of desert, or it might be a claim to an entitlement, or legitimate expectation. If it is meant to be a claim of desert, then that claim is at best undetermined by Rawls' theory, for as we have seen, the argument from arbitrariness works only to undermine individual argument from desert, not necessarily to install a social one. Moreover, for the desert, not necessarily to install a social one. Moreover, for the community as a whole to deserve the natural assets in its province and the benefits that flow from them, it is necessary to assume that society has some pre-institutional status that individuals lack, for only in this way could the community be said to possess its assets in the strong, constitutive sense of possession necessary to a desert base. But such a view would run counter to Rawls' individualistic assumptions, and in particular to his view that society is not 'an organic whole with a life of its own distinct from and superior to that of all its members in their relations with one another' (264).

The alternative would be to view society's claim on the distribution of assets as an entitlement given by legitimate expectations founded upon social institutions 'established for prior and independent social ends' (313), prior and independent, that is, to the entitlement itself. Rawls might argue along these lines that the difference principle does not assume that society has some pre-institutional

status denied to individuals, only that the parties to the original position would agree to regard the distribution of natural talents as a common asset and to share in the benefits of this distribution. The notion of society as the owner of natural assets for which individuals are the guardians would then be seen as the result of the original position and not its premise.

But this still leaves us wondering about the 'prior and independent social ends', and the source of their priority and independence. Somehow the society's claim on the distribution of assets must make a difference before institutions get established. *Some* social claim on the distribution of assets would seem a necessary presupposition of the agreement in the original position, otherwise the parties would be deliberating about how to allocate shares that were not (yet) rightfully *theirs* to allocate. Nozick argues, along similar lines,

Do the people in the original position ever wonder whether *they* have the *right* to decide how everything is to be divided up? Perhaps they reason that since they are deciding this question, they must assume they are entitled to do so; and so particular people can't have particular entitlements to holdings (for then they wouldn't have the right to decide together on how all holdings are to be divided) (1974: 199n).

To this Rawls might reply that no antecedent social claim is involved, since the parties to the original position are not faced, strictly speaking, with the moral question of how they *ought* to allocate distributive shares (which would indeed imply their having an antecedent 'right' to decide), but only with the prudential question of how, given the relevant constraints on information, etc., they would, from the standpoint of individual self-interest, *prefer* that individual shares be allocated. The original position, Rawls might remind us, is not an actual site of allocation, but rather a way of thinking.

Still, we want to know *why this* way of thinking is appropriate to questions of distributive justice, and whether the appropriateness of this way of thinking does not itself depend on the parties to the original position having some antecedent claim on the totality of natural assets. For even if the parties to the original position would, as a matter of rational prudence, reason in the way Rawls says they would, it is not immediately clear, absent some prior social claim, why their collective choice should determine the just distribution of

these assets. This leads us to the issue of justification and the question of how the argument from the original position serves to justify the principles that result. If society's claim to the distribution of natural assets can be shown to be a product of the original agreement rather than its premise, then Rawls will have resolved the stand-off in favor of the difference principle without having to attribute an antecedent claim of desert to society as a whole. If, on the other hand, the notion of common assets should turn out to be a presupposition of the original agreement, then Rawls will have overcome the stand-off only by implicit reliance on a claim of social desert, and hence reliance on a wider subject of possession, presumably the community, held to own the assets we individually bear. Once again an intersubjective dimension would intrude on Rawls' individualistic project. Although such a result would do considerable damage to the deontological ethic Rawls seeks to defend, I shall try to show that his version of contract theory cannot avoid it.