Neutralizing luck, rewarding effort[*]

Marc Fleurbaey[**]

July 2004

Introduction

In *Justice, Luck, and Knowledge*, Susan Hurley does a great service to the theory of egalitarianism by doing what most authors have shied away from doing so far, namely, opening the black box of "responsibility" in order to examine how the various conceptions of responsibility can inform the debate about the just allocation of resources in an egalitarian society. She ends up arguing in favor of disconnecting social responsibility (or accountability) from moral responsibility, so that moral responsibility need not determine exactly what is equalized in society, although, in her approach, it can still play some role in constraining the degree of redistribution (via incentives and its effect on well-being).

I am very sympathetic to both moves,[1] and found a richness of useful information and arguments in this book. In this short paper, I will rather focus on another argument that is central to the book and which I found quite striking and surprising. A key thesis put forth by Hurley is that the principle of neutralizing luck, or, more precisely, the effects of luck on distribution,[2] cannot serve as a justification for any egalitarian claim, nor even help determining the desirable degree of distribution. As she puts it, this principle can tell us *what* to distribute (i.e. welfare or resources or opportunities) but not *how* to distribute it (i.e. equally or otherwise).

Hurley's argument

Hurley's target of criticism is the idea that neutralizing luck is a fundamental egalitarian aim (Cohen 1989), so that it automatically entails a good deal of egalitarianism –with different degrees, admittedly, depending on how the scope of luck in people's fate is delineated. Her basic argument against this quite intuitive idea is as follows: 'If we equalize… goods that are a matter of luck, then there will be equalities between people with respect to goods that are a matter of luck. But people are no more responsible for an equal distribution of goods that are a matter of luck than they are for an unequal distribution.' (p. 151) She describes the '*egalitarian fallacy*' as the mistaken inference from (A) "It is a matter of luck that *a* and *b* are unequal" to (B) "It would not be a matter of luck if *a* and *b* where equal". (p. 152) She claims that we can obtain egalitarian conclusions only by assuming that equality is the default position and that responsibility can then be used only to justify departures from equality. But, obviously, in such an "equality-default" approach luck neutralization plays no independent role in obtaining egalitarian conclusions.[3] As a further illustration of this point, she imagines the possibility of an alternative view which takes *inequality* instead of equality as the default position: Under this alternative view, "aristocrats should have more than peasants, whether this is a matter of luck or not. Departures from this inequality, including equality, need to be justified by responsibility. Only equalities for which people are responsible are permitted, not equalities that are a matter of luck.' (p. 154)

That neutralizing luck does not provide an independent argument in favor of egalitarianism is not exactly a worry for egalitarians who take equality (or priority) as their basic principle anyway and are only in search of a precise specification of it. But to say that neutralizing luck cannot even help in determining how to redistribute the goods for which people are not responsible is quite striking and provocative. This double thesis plays a key role in various parts of Hurley's book, and in particular motivates her quest for an alternative basis for egalitarianism, which she finds in the principle of neutralizing bias (rather than luck) through veil-of-ignorance devices.

---

[3] There may be good reasons, however, to take equality as the default position. If we were all identical, life would be simple (and dull) and equality would trivially be achieved. We need to relate inequalities to relevant differences, even when we are not egalitarian.

Intuition restated

Hurley's argument is surprising because the luck egalitarian view that neutralizing luck means suppressing inequalities due to luck is very intuitive. Here I will try to restate the intuitive luck egalitarian view in order to see if and where the fallacy operates.

Suppose we decide to focus on a particular dimension of achievement in people's lives, say welfare or some notion of advantage. Suppose also that we agree on a separation between individual characteristics which are a matter of luck or a matter of responsibility. It is then a matter of fact, not of assumption or ethical principle, that, in any particular social situation, every individual's level of achievement depends on her characteristics of luck and responsibility. Formally, there is a function which determines every individual's achievement as a function of her characteristics, as in the following formula:

$$\text{achievement} = f(\text{personal luck, personal responsibility}).$$

The function f is itself determined by the laws of nature and the social institutions and policies which prevail in the contemplated situation. There is a lot of "luck" in the function f, but this is not the luck that luck neutralizers want to nullify. Of course, if the function f is unsatisfactory because the government is inefficient, then something should be done about it, but this is not directly a problem of distribution. What luck neutralizers would like to do is to remove "personal luck" from the arguments of f. This may be possible because f is partly determined by social policies. If we can produce a social situation and a corresponding function f^* such that the above formula can be written

$$\text{achievement} = f^*(\text{personal responsibility}),$$

then indeed we can say that the "effects of luck on distribution" have been "neutralized". This neutralization makes personal luck disappear from the relevant factors of personal achievement.

I hope this is a faithful description of the intuitive notion of luck neutralization that is pervasive in the egalitarian literature. There may be alternative ways,[4] and I will examine one of them later on. Now, the first question to ask is whether this operation of neutralization has any egalitarian consequences. The answer is clearly positive. Indeed, the second formula above has the following implication. Any pair of individuals who have the same characteristics of personal responsibility will obtain the same level of achievement,

---

[4] Some variants of resource-egalitarianism (Dworkin 2000, van Parijs 1995) would probably require a milder formulation for luck neutralization, but I will not examine this here, although this has to do with the conflict between luck neutralization and neutral reward alluded to below. See e.g. van Parijs (1997), Fleurbaey (2002) for details.

independently of their personal luck. Departure from equality of achievement will be observed only in case of different characteristics of personal responsibility.

Where is the fallacy?

The way in which this egalitarian implication is obtained may be traceable to the following simple argument. To neutralize personal luck means reducing its impact on personal achievement to zero for everybody. And zero for everybody is indeed a kind of equality.

Does this way of reasoning fall prey to the fallacy identified by Hurley? Her arguments seem to revolve around the fact that, in the second formula above, the function f^* itself still contains a lot of luck, so that it is not clear that luck is more neutralized in this situation than in another, less egalitarian one. It is true that luck is pervasive in any social situation, egalitarian or otherwise, but it is also true that the principle of luck neutralization focuses on one particular kind of influence of luck, not all kinds. Namely, it wants to cancel the direct impact of personal luck on personal achievement.

Things are tricky here because even when one can write

$$\text{achievement} = f^*(\text{personal responsibility}),$$

it is still the case that personal luck exists and has some potential effect on personal achievement. But this effect is compensated by the neutralizing effect of social policy, so that the total effect is null. Therefore, one can agree with Hurley that in this situation it is still a matter of luck that personal achievement depends on personal characteristics in this particular way, but agree with luck neutralizers that it is only in this kind of situation that the undesirable effect of personal luck is neutralized. In conclusion, Hurley understands the phrase "effects of luck on distribution" in a broader way than luck neutralizers. She considers the idea of eliminating all effects of all kinds of luck on distribution, and can indeed make the easy case that this is essentially impossible and does not point out any particular distribution as more salient than others. In contrast, luck neutralizers do not care, qua luck neutralizers,[5] about the general effect of luck on our lives and focus only on the relation between *personal* luck and *personal* achievement for every individual.

Does the narrower sense adopted by luck neutralizers smuggle in some egalitarian premiss? The luck to be neutralized is "the impact of personal luck on personal achievement". It is possible to interpret this expression in terms of interpersonal comparisons and inequality reducing, so that neutralizing luck is immediately understood as removing inequalities due to differential personal luck. This is indeed the most frequent way of presenting the basic

"compensation principle" in the economic literature I am familar with.[6] Now, consider a purely personal, counterfactual interpretation of the above expression.[7] Neutralizing luck, in this sense, would mean that a counterfactual change in personal luck for a particular individual should not change her personal achievement.[8] In relation to this second interpretation, one may imagine the following situation, inspired by Hurley's example of aristocrats and peasants. The government might select an arbitrary distribution and make sure that for everybody, individual achievement is no longer influenced by personal luck, so that any change in personal luck triggers a strictly personalized compensating action from the government. Formally, such a situation would have

$$\text{achievement} = f^*(\text{personal responsibility}),$$

but now the function f^* would be specific to each individual, being arbitrarily set up by the government. In this case the beautiful equality of achievement for individuals with identical responsibility characteristics would no longer be obtained.

What would a luck neutralizer answer to this example? That it does not fit in the framework. Now the function f^* is itself a personal characteristic. Does it belong to personal luck or to personal responsibility? If it belongs to personal luck, then obviously the impact of luck has not been cancelled in this example, which is then inconsistent (in Hurley's example, what makes an individual an aristocrat or a peasant?). If it belongs to personal responsibility, then the above formula is ill-specified, and it is indeed the case that individuals with identical responsibility characteristics (including f^*) have equal levels of achievement.

In conclusion, it seems that even under a purely intra-personal, counterfactual understanding of the above definition of luck, it is the case that neutralizing luck has egalitarian implications.

Redistributing luck goods

There is another framework which may serve for the description of the luck neutralizing idea, and it is suggested by Hurley's frequent reference to "goods that are a matter of luck". In the above framework, luck and responsibility applied to personal characteristics which were left

---

[5] Of course, any reasonable theory advocates fighting bad luck in general, this is just a matter of efficiency.

[6] It is surveyed in Fleurbaey (1998) and Fleurbaey and Maniquet (forthcoming).

[7] The distinction between interpersonal and counterfactual interpretations is repeatedly made by Hurley, and this is quite useful.

[8] Nor should it change the achievement of other individuals, since this is presumably a matter of luck, a fortiori, for them too. It is in general impossible to achieve such a state of affairs. Personal luck characteristics may alter the set of feasible allocations in such a way that it is impossible to make all personal achievements impervious to changes in the distribution of personal luck characteristics. I ignore this practical impossibility for the sake of the argument.

unspecified otherwise, but which operated as factors in the analysis of some chosen measure of achievement. One can instead introduce the separation between luck and responsibility in the space of achievements as well as the space of characteristics. For simplicity, let us imagine that in some way of other, redistribution can operate over "goods" (which may be achievements or characteristics), and that, after goods have been sorted out by the luck-responsibility filter, the idea is to neutralize the effect of luck on the distribution of goods generally speaking. Does this have egalitarian implications or is it a fallacy?

This different framework amounts to generalizing the previous one by allowing achievement to be multidimensional. Personal achievement is now a bundle of goods, and in any situation it can be trivially described as related to personal luck (goods that are a matter of luck) and personal responsibility (goods that are a matter of responsibility). Now consider a situation in which this personal bundle is simply related to personal responsibility and not linked, in the description of the distribution (or, if one prefers, counterfactually), to personal luck. What can this mean, since the bundle still contains bundles which are a matter of luck? This simply means that we have a relation like this one:

$$(\text{luck goods, responsibility goods}) = f^*(\text{responsibility goods}),$$

and since the function f^* is the same for all individuals,[9] this logically entails that luck goods are equally distributed. Again, it seems hard to avoid the conclusion that some kind of equality must be enforced when the link between personal luck and personal achievement is nullified, even if one accepts Hurley's point that there is still a lot of luck everywhere.

Neutralization and reward

In conclusion, it is rather dubious that the idea of neutralizing luck has no egalitarian implications. Whether the goal of neutralizing luck provides a good ethical foundation for egalitarianism is of course not settled by this point, and I share Hurley's skepticism about giving moral responsibility such a massive role in the design of social institutions. But I remain unconvinced that neutralizing luck cannot serve to (badly) justify some kind of (questionable) equality. Certainly I am not convinced that it says nothing about *how* to redistribute (as opposed to *what*). The principle of neutralizing the impact of talents and handicaps has such an important role in the economic analysis of fair redistribution, and appears to be so constraining in the selection of allocations, that it is unimaginable to discard it as empty. Truly enough, it is often directly formulated in terms of equality of achievement

---

[9] If it were not identical across individuals, one would have to ask again whether it is a matter of luck or responsibility, and a similar reasoning as above would force us to accept that it must be equal, as some level.

for individuals with identical responsibility characteristics, and Hurley's challenge is quite valuable in forcing us to think more about the link between this "egalitarian" formulation and the "neutralization" or "compensation" language.

If there is indeterminacy in the principle of luck neutralization, it lies, I believe, elsewhere. Look again at this formula:

$$achievement = f^*(personal\ responsibility).$$

Even if it encapsulates the principle of luck neutralization correctly and is quite constraining over the distribution, it does not guarantee that there is only one way to neutralize luck. There may be several allocations of this kind, each with a different function f^*. In other words, neutralizing luck still leaves unanswered the question of how personal achievement should be related to personal responsibility, i.e. of how responsibility should be *rewarded*. Hurley identifies this indeterminacy (p. 160-162) but seems to make the wrong inference that indeterminacy on the reward issue somehow extends across the board to the neutralization issue.

She may be forgiven because the question of reward has been blurred in a good part of the literature. Many authors have focused on luck neutralization, while others have swiftly adopted a particular reward principle without much pondering or even without noticing. Hurley nonetheless alludes to the reward issue in various parts of the books. Here is an interesting quotation: '[In luck egalitarian theories,] we only aim to redistribute goods that are a matter of luck for people, not goods people are responsible for. For instance, we do not redistribute goods people are responsible for via their choices, including the outcomes of chosen gambles.' (p. 150) The first sentence is ambiguous, like many similar ones in the literature. Does it mean that responsibility goods should not be redistributed, or that nothing is said about whether they should or not? The second sentence, fortunately, clarifies the point and clearly suggests that the former is true, and I believe that this is indeed faithful to the main luck egalitarian theories. But then it is hard to claim at the same time that luck egalitarianism is indeterminate about reward. Saying that goods that are a matter of responsibility should not be redistributed provides a quite substantial indication about how responsibility and achievement should be related, i.e. about what it means to give people what they are responsible for.

It seems to me that the concept of responsibility is so closely linked to this idea of neutrality, i.e. of no intervention after individuals' responsible decisions, that it is indeed a little artificial to separate it from the principle of luck neutralization. But they are logically distinct. One can think of a system of perfectionist rewards in which, after personal luck has

been neutralized, particular behaviors are positively rewarded and others are punished by special intervention of public authorities. Such a system is compatible with luck neutralization and with some exercise of personal responsibility. In conclusion, it appears that a full-blooded theory of luck egalitarianism normally encompasses the idea of luck neutralizing and the idea of neutrality with respect to responsible decisions, as in the above quotation. But these two ideas are distinct, and can be separated. Luck neutralization in and of itself does not give any indication about reward.

Roemer's reward scheme

Are there luck egalitarian theories which do not retain this idea of neutrality? Roemer's approach is an example. But I disagree with Hurley about the reasons why it exemplifies this kind of theory. According to her, 'giving people what they deserve on account of their efforts is not the same as giving them what they are responsible for' (p. 182), and Roemer's redistribution scheme does the former but not the latter, in a perfectionist way. She finds perfectionism (or 'paternalism') in Roemer's scheme because it compares the exercise of responsibility by individuals according to their relative level of effort or achievement, each in their own type[10]. This comparison itself, for matters of redistribution from advantaged types to disadvantaged types, is, according to her, based on a prior selection of the direction of effort that we want to reward.

This is actually mistaken, I think, because there is nothing in Roemer's scheme that depends on a positive or negative relation between effort and personal achievement. Roemer's scheme would apply exactly in the same manner under any of these configurations, in the sense that it would have the same distributive implications independently of whether effort enhances personal achievement or dampens it. This scheme is not geared toward rewarding effort, contrary to what Hurley claims.

But it is not neutral with respect to responsible decisions, and this can be most easily understood when looking at the special case in which there is only one type, that is, when personal luck is already the same for everybody, prior to any redistribution. In such a case the neutrality ideal advocates the absence of redistribution.[11] In contrast, Roemer's scheme then boils down to maximizing the sum total of personal achievements and is thus similar to classical utilitarianism (applied not to utility but to the considered achievement). It therefore advocates redistribution in favor of those with responsibility characteristics which give them a

---

[10] A "type" here is, as defined in Roemer (1998), a group of people with the same luck characteristics.

higher productivity in the transformation of resources into achievement. This scheme does not reward effort but responsibility decisions which entail a high marginal rate of transformation of resources into achievement. This is not perfectionist in the way Hurley sees it.[12]

One possible justification for Roemer's utilitarian kind of reward is the following. One may interpret the idea that we only want to redress inequalities due to luck as meaning that our social criterion should display aversion to inequality only in the direction of luck. In the direction of responsibility, a zero aversion to inequality is supposed to reflect the idea that we are indifferent to inequalities due to personal responsibility. In welfare economics, a social welfare function with a zero aversion to inequality is computed as the sum of individual utilities. Accordingly, Roemer's criterion applies the maximin in the direction of luck, and the sum in the direction of responsibility.

Oppositions and dilemmas

Roemer's criterion is the only one discussed in Hurley's book, but there are other relevant and possibly more attractive criteria which adopt the same utilitarian kind of reward or a neutral kind of reward. The opposition between these two kinds of reward has not received enough attention among philosophers, and reflects, in the field of luck egalitarianism, the deep classical opposition between utilitarianism and libertarianism. When people are responsible, should we intervene in order to maximize the total outcome or let them in peace? That is, in a nutshell, the question.

Here is an example of another criterion, which has been proposed by Van de gaer (1993). Like Roemer's scheme, it does apply the maximin for luck and the sum for responsibility, so that it also adopts the utilitarian conception of reward. But it performs these two operations in the reverse order. Roemer first computes the minimum across types for all values of responsibility characteristics, and then sums up these figures. Van de gaer first computes the mean outcome for each type, and then applies the maximin criterion to these means. I present this criterion here because it shows that Roemer's complex comparison of the degree of responsibility for people in different types is not so essential. With Van de gaer's criterion, there is no need for such a comparison. It is very easy to compute the mean outcome per type, and in this operation there is no need to evaluate comparable degrees of effort or

---

[11] This is not the same as "laisser-faire" in general, since market failures may warrant state intervention for efficiency purposes. But redistribution is not needed.

[12] Paternalism appears, however, when the scheme is applied to a special dimension of achievement, such as education, instead of general individual well-being. Achieving equal opportunity for education with Roemer's criterion may lead to policies which are Pareto-inefficient because people do not care only about education.

responsibility across types. Van de gaer's criterion is, for practical purposes, quite similar to Roemer's (although there are some interesting differences).

The non-liberal feature of both criteria is striking in some applications, and here is a canonical example, which I expound here in order to emphasize the need for a careful evaluation of reward schemes. Suppose that personal achievement, personal luck and personal responsibility are all measured by real numbers, and that personal achievement is determined by the following formula:

$$\text{achievement} = (\text{money} + \text{personal luck}) \ast \text{personal responsibility}.$$

This formula is quite typical of a situation in which your total resource endowment depends on some inherited part (personal luck) and the prevailing policy (money transfers), and your achievement depends on how prudently you use your total resource. Suppose that money can be redistributed without cost and limit by the government (and without affecting personal characteristics, including responsibility characteristics). Consider a population with two types, the lucky and the unlucky. Their personal luck equals, respectively, 3 and 1. Each type is subdivided in two equal-sized subgroups with different values for the responsibility characteristic, 3 and 1 again.

A natural, liberal policy[13] in this case, is to give two additional units of money to the unlucky in order to compensate for their bad luck, and to give these two units to all of them, independently of their personal responsibility. This gives the following values of achievement, when the amount of available money, per capita, is 4 (so that the unlucky receive 5, and the lucky 3 each):

	responsibility = 1	responsibility = 3
luck = 1	6 = (5+1) * 1	18 = (5+1) * 3
luck = 3	6 = (3+3) * 1	18 = (3+3) * 3

Table 1

But this is not the optimal policy for Roemer's and Van de gaer's criteria, which both advocate policies which give money only to those with a high value for the responsibility parameter, because they have a higher marginal rate of transformation of money into achievement. For Roemer's criterion, this yields the following distribution of achievement:

[13] It is advocated by all known criteria based on the neutral kind of reward. One such criterion equalizes the counterfactual outcome that individuals would have if their personal responsibility corresponded to some norm or reference value. Another criterion makes every individual have the same outcome as she would have, counterfactually, with some reference luck and her actual personal responsibility characteristics. Without referring to the latter criterion, Hurley objects to this kind of counterfactual analysis, arguing that luck and responsibility are 'nonseparable' (p. 168), meaning that luck may have an influence over responsibility. This

	responsibility = 1	responsibility = 3
luck = 1	1 = (0+1) * 1	30 = (9+1) * 3
luck = 3	3 = (0+3) * 1	30 = (7+3) * 3

Table 2

This example shows that the utilitarian kind of reward may lead to strong redistribution within types, in order to favor those who have favorable sorts of responsibility characteristics. Moreover, this drive toward rewarding high "marginal utility" may run into conflict with the goal of neutralizing luck, as illustrated in the first column of Table 2, in which the unlucky with low responsibility value fail to be compensated for their bad luck.[14]

This is also a general result that comes out of the economic literature on this topic. There is a tension between luck neutralization and each of the two kinds of reward, neutral and utilitarian. Full luck neutralization often leads to transfers that violate the standard of reward.[15]

Conclusion

In brief, my criticism is that Hurley may be too pessimistic about the ability of luck egalitarianism to give precise indications about the just distribution, and a little superficial about the issue of reward, like a good part of the literature. As usual, this paper focuses on criticism because this is how its author thinks he can contribute to pushing the debate forward. But this does not impugn Hurley's great merit in connecting the literature on moral responsibility to the literature on distributive justice.[16] That she ends up concluding that this connection is not as useful as imagined by luck egalitarians is quite interesting and will encourage those who intuitively shared this view and looked for a different approach to the division of labor between society and individuals in just institutions. This, however, leads me

---

would deserve a longer discussion, but suffice it to say that Hurley's argument does not show that it is irrelevant to compare the fates of individuals with similar personal responsibility and different personal luck.

[14] Van de gaer's criterion leads to a slightly different policy, which interestingly overcompensates in the second column:

	responsibility = 1	responsibility = 3
luck = 1	1 = (0+1) * 1	31 = (9.33+1) * 3
luck = 3	3 = (0+3) * 1	29 = (6.66+3) * 3

Table 3

See Fleurbaey and Maniquet (forthcoming) for a more detailed discussion of this and similar examples.

[15] The conflict between luck neutralization and neutral reward does not appear in the above example (Table 1), but occurs when achievement is related to personal characteristics in such a way that money transfers required to neutralize luck differentials have different values for different responsibility characteristics. For instance, if the hard-working unlucky need more transfers than the lazy unlucky, luck neutralization implies some special (non-neutral) reward for hard work.

[16] Moreover, she provides every non-specialist of either field (or both) with a useful overview of the 'philosophical landscape'.

to a last question. As an alternative to luck egalitarianism, Hurley proposes to equalize a list of objective goods (food, shelter, clothing, etc.). The list presumably will not contain all objective goods (like e.g. understanding Einstein's relativity), and, she explains, 'in some cases, the… good is the end-state, such as being adequately nourished, while in others it is the opportunity or the access, as in family life' (p. 248). Really, can this kind of issue (selecting the goods, defining opportunities) be addressed without giving responsibility (not necessarily moral responsibility) a central role in the definition of *what* to distribute?

References

Cohen G.A. 1989, "On the currency of egalitarian justice", *Ethics* 99: 906-944.

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

Fleurbaey M. 1995, "Equal opportunity or equal social outcome?", *Economics & Philosophy* 11: 25-55.

Fleurbaey M. 1998, "Equality among responsible individuals", in J. F. Laslier et alii eds, *Freedom in economics*, Routledge.

Fleurbaey M. 2001, "Egalitarian opportunities", *Law and Philosophy* 20: 499-530.

Fleurbaey M. 2002, "Equality of resources revisited", *Ethics* 113: 82-105.

Fleurbaey M., F. Maniquet (forthcoming), "Compensation and responsibility", in K.J. Arrow, A.K. Sen, K. Suzumura eds., *Handbook of Social Choice and Welfare*, Amsterdam: North-Holland.

Hurley S.L. 2003, *Justice, Luck, and Knowledge*, Cambridge, Mass.: Harvard University Press.

Roemer J. E. 1998, *Equality of opportunity*, Cambridge, Mass.: Harvard University Press.

Van de gaer D. 1993, *Equality of Opportunity and Investment in Human Capital*, Ph.D thesis, K. U. Leuven.

van Parijs P. 1995, Real *Freedom for All. What (if Anything) Can Justify Capitalism?*, Oxford: Oxford University Press.

van Parijs P. 1997, "Social justice as real freedom for all: A reply to Arneson, Fleurbaey, Melnyk and Selznick", *The Good Society* 7(1): 42-48.