

MICHAEL HUEMER

## NON-EGALITARIANISM

### 1. INTRODUCTION

The general idea of egalitarianism is that equality, in some sense, is a value.<sup>1</sup> *Inegalitarianism* (a view which few if any hold) would be the view that equality is a disvalue or inequality a value. By contrast, *non-egalitarianism* holds that equality is neither good nor bad, but morally neutral. There are multiple possible forms of egalitarianism (and, correspondingly, of inegalitarianism and non-egalitarianism), depending both on how one interprets “equality” and on how one interprets “is a value”. In the following, I shall be concerned with one particular form of egalitarianism, which I call *Axiological Welfare Egalitarianism*.<sup>2</sup>

For present purposes, I shall treat as primitive the notion of an individual’s *level of well-being* at a given time. You may, if you wish, understand this in terms of pleasure, in terms of preference-satisfaction, or in terms of some notion of objective interests. I intend my following arguments to apply regardless of your theory of well-being, with just one restriction: I assume that level of well-being is a quantity, such that it makes sense to speak of one person being twice as well-off as another, or 1.5 times as well-off, etc.

I shall use the term “event” to include any (nonzero) portion of the life of an individual; I also assume that fusions of events are events.<sup>3</sup>

For a given portion of an individual’s life, I define the *utility* the individual enjoys in that portion of his life (the utility of that event) as the integral of his level of well-being over the relevant time interval. I define the *total utility* of an event as the sum of the utilities of all the individuals involved in the event (i.e., the sum of the utilities, for each individual a portion of whose life is included in the event, that those individuals enjoy during the relevant portions of their lives). Thus, total utility, by definition, is additive across persons: if Sue gets  $U_S$  units of utility during her life and Liz gets



$U_L$  units of utility during her life, then the total utility of the event that is the fusion of Sue's life and Liz's life is  $U_S + U_L$ .

The total utility of an event is not to be confused with the *utility-value* of the event. The utility-value of an event is the amount of non-instrumental *good* that the event contains in virtue of the allocation of well-being among individuals involved in the event. Utility-value, by definition, includes any value arising from the total utility of an event, as well as any value or disvalue arising from the manner in which that utility is distributed. A useful way to think about it is that it is the amount of value one could infer that the event has, if one is given a description of each person's level of well-being at all times when that person is involved in the event and one knows that there are no other morally relevant factors.

Well-being, individual utility, and total utility are all quantities. I believe utility-value is also a quantity, though my argument requires no more than that there be a total ordering in terms of utility-value.

We can now define the following three ethical theses:

*Axiological Welfare Egalitarianism:* Equality in the distribution of utility among persons is intrinsically (non-instrumentally) good, and/or inequality in such distribution is intrinsically bad.<sup>4</sup> We will focus on the following logical consequence of this thesis: given two possible histories of the world with the same total utility, the history (if any) in which utility is more evenly distributed has more utility-value.<sup>5</sup>

*(Axiological Welfare) Inegalitarianism:* Inequality in the distribution of utility among persons is intrinsically good, and/or equality in such distribution is intrinsically bad. This entails that, given two possible world-histories with the same total utility, the history (if any) in which utility is more evenly distributed has less utility-value.

*(Axiological Welfare) Non-Egalitarianism:* Both equality and inequality in the distribution of utility among persons are intrinsically indifferent. Thus, two world-histories with the same total utility have the same utility-value, regardless of the evenness of the distribution.

## 2. AN ARGUMENT FOR NON-EGALITARIANISM

The argument for Non-Egalitarianism uses three general moral principles:

*Intrapersonal Non-Egalitarianism:* Equality in the distribution of utility across times in a single individual's life is evaluatively neutral.

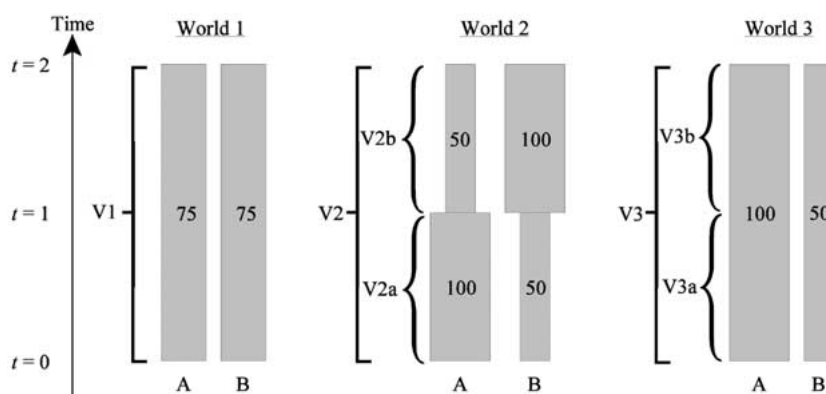
*Strong Supervenience of Utility-Value:* If two possible events are intrinsically identical in terms of non-evaluative properties, then they have the same utility-value.<sup>6</sup>

Here is a more precise, formal statement of the idea, relying on Kim's conception of events:<sup>7</sup> Suppose that  $n$  people exist during some time interval  $T_1$ . Let  $R^n$  be an  $n$ -place predicate giving a possible complete description of the goings-on in the lives of these  $n$  people during  $T_1$  (including their relations with each other). Let  $\langle a_1, \dots, a_n \rangle$ , be an ordering of the  $n$  people. Let  $\langle b_1, \dots, b_n \rangle$ , be a permutation of this ordering. And let  $T_2$  be some time interval of equal length to  $T_1$ . Then the utility-value of the event  $[\langle a_1, \dots, a_n \rangle, R^n, T_1]$  equals the utility-value of the event  $[\langle b_1, \dots, b_n \rangle, R^n, T_2]$ .

*Cross-Temporal Additivity of Utility-Value:* Given two events, A and B, occupying nonoverlapping time intervals, the utility-value of the fusion of A and B equals the utility-value of A plus the utility-value of B.

Now consider the following three possible worlds (see Figure 1). Each world contains two people, A and B. In the first world, A and B each have a level of wellbeing of 75 throughout their lives. V1 is the utility-value of the history of world 1.

In the second world, A spends the first half of his life at a higher level of wellbeing, 100, and the later half at 50. B spends the first half of his life at 50, and the second half at 100. V2a is the utility-value of the first half of the world's history; that is, the utility-value of the event that is the fusion of the first half of A's life and the first half of B's life. Similarly, V2b is the utility-value of the second half of world 2's history. V2 is the utility-value of world 2 as a whole.



*Figure 1.* The vertical dimension on the page represents time. The width of the bars indicates the level of well-being that each individual enjoys; the height of the bars indicates duration. Total utility enjoyed during a period of an individual's life is the area of the bar representing that period.

In the third world, A enjoys a level of well-being of 100 throughout his life, while B enjoys a level of 50 throughout his life. V3 is the utility-value of this world, V3a the utility-value of the first half of this history, and V3b the utility-value of the later half.

All three worlds have the same total utility. World 1 has an equal distribution of this utility, while world 3 has a highly unequal distribution. I shall argue that worlds 1 and 3 have equal utility-value:

1.  $V1 = V2$       Given Intrapersonal Non-Egalitarianism, the utility-value of A's life is identical in the two worlds, as is the utility-value of B's life, and both worlds enjoy perfect interpersonal equality.
2.  $V2a = V3a$       From Strong Supervenience of Utility-Value.
3.  $V3a = V3b$       From Strong Supervenience.
4.  $V2a = V2b$       From Strong Supervenience.
5.  $V2b = V3b$       From 2, 3, 4.
6.  $V2a + V2b = V3a + V3b$       From 2, 5.
7.  $V2 = V2a + V2b$       From Cross-Temporal Additivity.
8.  $V3 = V3a + V3b$       From Cross-Temporal Additivity.
9.  $V2 = V3$       From 6, 7, 8.
10.  $V1 = V3$       From 1, 9.

Axiological Welfare Egalitarianism entails that world 1 is better than world 3. Thus, the above argument, if sound, refutes Egalitarianism and establishes Non-Egalitarianism.

### 3. THE PREMISES EXAMINED

#### 3.1. *Are Worlds 1 and 2 Equally Good?*

It seems that  $V_1 = V_2$ , since neither A nor B would have reason to prefer either world over the other, and A and B are the only people who exist in either world. Furthermore, the two worlds have the same total utility, and both have a perfectly even distribution of that utility across persons.<sup>8</sup> How might one resist this?

*First:* one might want to resist (1) by saying that one has some degree of time preference (one prefers to get one's utility sooner rather than later), so that, if given a choice between A's life in world 2 and B's life in world 2, one would prefer A's life; similarly, if given a choice between B's life in world 1 and B's life in world 2, one would prefer the former. This might motivate a rejection, both of the claim that A and B are each as well off in world 2 as they are in world 1, and of the claim that world 2 is egalitarian.

However, while most people exhibit some degree of time preference, there seems to be no rational justification – and in particular, no justification in terms of the total amount of good obtained – for this preference.<sup>9</sup> There does not seem to be any reason for holding that goods that happen in one year are *ceteris paribus* more important (contribute more to the total good) in the history of the world than goods that happen in the following year.<sup>10</sup> (Would we also hold that past events are *ceteris paribus* more important than present events?)

It may be argued that some persons gain more total pleasure from having their (otherwise identical) pleasures sooner rather than later, perhaps because, due to their impatience, they suffer some during the waiting. Other persons may gain more total pleasure from deferring (otherwise identical) pleasures, perhaps because they enjoy the anticipation. However, neither of these possibilities is relevant here. For the *total utilities* in my hypothetical worlds are *stipulated*, and they are stipulated to be such that A's total utility over his lifetime in world 1 = B's total utility in world 1 = A's total utility in world 2 =

B's total utility in world 2. Thus, if world 1 lacks pleasures of anticipation (or pains of impatience), then either world 2 also lacks such pleasures (or pains), or they are exactly made up for in some other form of utility (or disutility).

The general point here is that our question is not whether equality of certain goods is *instrumentally* good – for example, whether it causally contributes to total utility. Our question is whether equality of welfare is *intrinsically* good – that is, whether it contributes to utility-value when total utility is fixed. Hence, we wish to consider situations in which total utility *is* fixed; we shall not entertain arguments to the conclusion that there is less total utility in one of our hypothetical worlds than there at first appears.<sup>11</sup>

*Second:* Perhaps the Egalitarian would object to step 1 by extending his egalitarianism to the intrapersonal level, holding that A's life is worse in world 2 than in world 1, due to the inequality between A's youth and A's old age, and similarly for B. This would mean that, if one has a choice between a life of constant, moderate well-being and a life of 'ups and downs' – varying levels of well-being ranging from low to high – with the same average for the two lives, one should prefer the former. No doubt some people would express a preference for the former, while others would express a preference for the latter. Either preference strikes me as merely arbitrary.

Notice that Intrapersonal Non-Egalitarianism is equivalent to the idea that it is prudentially rational to sacrifice some utility now as long as one thereby gains more utility later, or to sacrifice some utility later as long as one thereby gains more utility now.<sup>12</sup> The *intrapersonal* egalitarian would have to deny this principle. That strikes me as a good deal more counter-intuitive than the *interpersonal* Egalitarian's denial that it is always desirable to sacrifice one person's utility to gain a greater increment of utility for *someone else*. One might be tempted to think that A is worse off in world 2 than in world 1 because, for example, in his old age, he looks back upon how much more fun he had in his youth, and this makes him feel disappointment. But again, this sort of supposition is irrelevant, as the total utilities in the worlds are stipulated; by stipulation, either A does *not* feel any such disappointment, or it

is exactly compensated for by an increase in some other form of well-being.

*Third:* One might object to the claim that world 2 is egalitarian by challenging the assumption (call it the “complete lives view”) that the amount of (interpersonal) inequality in a world should be computed by comparing *complete lives*. If we make this assumption, we find world 2 to be interpersonally egalitarian, since A’s utility over A’s complete life equals B’s utility over B’s complete life. But an alternative way of assessing interpersonal inequality (call it the “simultaneous segments view”) would be to compare simultaneous time-slices of individuals – i.e., to first determine, at *each time*, the welfare inequality between A and B at that time, and then to integrate these values over the time interval from  $t = 0$  to  $t = 2$ . The simultaneous segments view yields the result that world 2 is highly inequalitarian, since there is a high degree of interpersonal inequality at every time during its history.<sup>13</sup>

At worst, this objection might require me to weaken my thesis to a refutation of *complete lives* Egalitarianism; however, I do not think even this qualification is necessary. Temkin and McKerlie both criticize the complete lives view by offering alleged counter-examples: essentially, they describe cases like world 2 and then invite us to appreciate intuitively that there is something bad about them in terms of inequality. Unfortunately, this does not strike me as obvious at all.<sup>14</sup> Rather, the complete lives view strikes me as the most natural one for an Egalitarian to take, and I find it intuitive that world 2 is interpersonally egalitarian. Moreover, there is a simple argument that world 2 is interpersonally egalitarian: if you are going to be dropped into a world like world 2, you have no reason, from the point of view of self-interest, to prefer either A’s or B’s position in it; setting aside time preference, you should be indifferent between the two possible fates. This makes it very difficult to claim that there is an interpersonal inequality in world 2. To amplify the point: Egalitarianism is usually motivated by the idea that inequality involves *unfairness*.<sup>15</sup> But *to whom* might world 2 be unfair? It would be arbitrary to choose either A or B. And since the complaint of inequality is a complaint regarding how a person or group fares *relative to others*,<sup>16</sup> it seems likewise impossible to

maintain that *both* A and B – that is, everyone in the world – are the victims of the unfairness.

The simultaneous segments Egalitarian will want to answer that the *first* half of the world is unfair to B, while the *second* half is unfair to A. But this does not answer the question; it does not tell us to whom he thinks *the world as a whole* is unfair. From the (putative) unfairness of both the first half and the second half of world 2, we could infer that the world as a whole is unfair only if we assume that unfairness is additive over time. But, *if* each half of world 2 contains some unfairness, then world 2 seems to be a counterexample to precisely that assumption. The difficulty of saying to whom the world as a whole is unfair, despite the ease of saying to whom each half is unfair, gives us grounds to reject the additivity of unfairness (or to reject that either half is unfair).

The simultaneous segments view has two further problems. First, it implies that world 2 is just as bad, in terms of inequality, as world 3 is. I think that very few Egalitarians would accept this. Second, suppose you know that I have a dentist's appointment next Friday at 10:30 (assume such appointments are unpleasant and thus bring about subnormal welfare for their duration). You have a dentist's appointment (with a different dentist) scheduled for 11:00 that day. Other things equal, should you change your appointment to 10:30, so you and I will be suffering simultaneously? This would seem quite arbitrary, yet the simultaneous segments view would say you have good reason to do so.<sup>17</sup> Combining these two points, the simultaneous segments view would say that the reason one has for *synchronizing* the suffering of two people is just as strong as the reason one has, if one of a pair of people is suffering, for distributing half of the one person's suffering to the other – since both actions eliminate an inequality (and one can describe the cases such that the same magnitude of inequality is involved).<sup>18</sup>

### 3.2. *The Value of Compensation*

Now turn to premise 2. Whether or not one accepts the Strong Supervenience of Utility-Value *in general*, it would seem bizarre to hold that the first half of world 2 was better or worse than the first half of world 3, given that they are identical. It would be equally strange to deny premise 3. Why would one half of world 3 be better than



the other (identical) half? Perhaps the Egalitarian would argue that the goodness of a given segment of a world's history is a relational property, depending in part on what goes on in the rest of the world's history. Thus, one might argue, against premise 4, that the later half of world 2 is better than the earlier half, because in the later half, B is 'compensated' for his relative misfortune in the earlier half, and this adds some value to the event (or the absence of such compensation subtracts from the value of the earlier half of the world's history, or both). Note that this objection is distinct from the objection to follow in which Cross-Temporal Additivity is denied; the present objection allows that the value of world 2 is the sum of the value of its first half and the value of its second half but maintains that the two halves have different value. This is to deny the strong supervenience of utility-value, though one could still maintain *global* supervenience (if two worlds have qualitatively identical *total histories*, then they have the same utility-value).<sup>19</sup>

But on reflection, it seems arbitrary to choose either half of world 2 as being better than the other. May not the earlier half of the world be said to compensate for the later half, as easily as the later half may be said to compensate for the earlier? Some would say that the notion of compensation is inherently backward-looking: one can only compensate for what has already happened, not for future events. But suppose I own a machine that sometimes produces unpleasant noise, which I have to compensate nearby residents for. Usually, I pay them compensation after the fact. Today, however, I know that there will be a lot of noise tomorrow, so I pay the residents in advance. When tomorrow comes, do I have to pay them again, because the earlier payment didn't count as compensation? Admittedly, this is a somewhat different kind of 'compensation' from that allegedly involved in world 2. But if the one kind of compensation need not be backward-looking, why should the other?<sup>20</sup>

Furthermore, if world 1 is to be as good as world 2, and one half of world 2 better than the other half, then one half of world 2 (either the earlier or the later half) must be superior to half of world 1. If we say that  $V_{2b} > V_{2a}$ , then we must hold that the second half of world 2 is better than the second half of world 1. And it seems odd, from an egalitarian perspective, that an event containing a large inequality, albeit one that follows another inequality in the opposite

direction, should be considered superior to an event that continues a perfectly equal distribution of utility. This sounds suspiciously like the doctrine that “two wrongs make a right” – more precisely, that a second *prima facie* bad is superior to a second good.

### 3.3. *Cross-Temporal Additivity*

An initially more promising approach would be to allow that each half of world 2 is as good as the other (and similarly for world 3) but to deny Cross-Temporal Additivity. Egalitarians, after all, already deny the analogous principle of additivity across *persons*. One might hold (i) that the overall utility-value of world 2 is *greater* than  $V2a + V2b$  because world 2 gains extra value from the fact that each half of its history ‘makes up for’ the inequality existing in the other half, and/or (ii) that the overall utility-value of world 3 is *less* than  $V3a + V3b$  because world 3 has the negative trait of an uncompensated-for inequality.

Indeed, it may seem *obvious* that an Egalitarian would deny Cross-Temporal Additivity, for the following reason:

- (a) Inequality is not additive over time.
- (b) Inequality partly determines value.
- (c) Therefore, value is not additive over time.

Premise (b) is the Egalitarian’s central thesis. Premise (a) is demonstrated by the example of our world 2, once we have rejected the simultaneous-segments method of measuring inequality: the first half of world 2 is highly unequal between persons, the second half of world 2 is highly unequal between persons, but world 2 as a whole is *equal* between persons. Thus, it seems that the Egalitarian would simply regard our example as refuting Temporal Additivity.

I think (a) is clearly true, and (c) is extremely plausible given (a) and (b). So I think this is the most natural position for an Egalitarian to take. Furthermore, the Egalitarian might charge me with begging the question in asserting Temporal Additivity, given how closely non-additivity is thus connected with Egalitarianism – though I might likewise charge him with begging the question in giving this reason for rejecting Temporal Additivity. But I shall not enter into that particular debate now. What we should recognize at this point is that the Egalitarian who is drawn to this position should also be

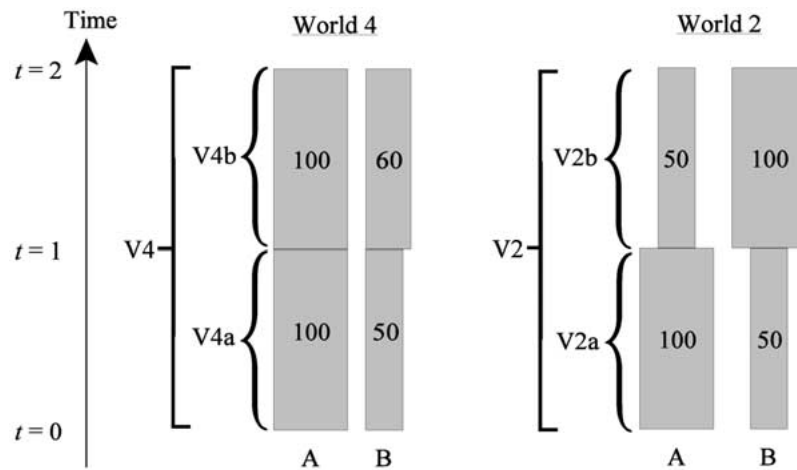


Figure 2. World 4 contains more total utility than world 2, but not enough more to compensate for the interpersonal inequality in world 4.

prepared to concede that, if strong, independent grounds were to be given for accepting Temporal Additivity, then he ought to give up Egalitarianism.

Such grounds can be given. The denial of Cross-Temporal Additivity generates counter-intuitive decision-theoretic results. Assume that world 3 is worse than world 2 on Egalitarian grounds. The Egalitarian will presumably agree that it is possible to add some amount to B's well-being in the second half of world 3 while still leaving world 3 inferior to world 2. This must be true unless the value of equality is *negligible* in comparison to that of utility. Suppose, then, that adding 10 units (the particular number doesn't matter) would accomplish this – that is, world 4 (see Figure 2) is still worse than world 2, though, obviously, better than world 3.

Now imagine two simple choice problems, and assume in each case that there are no relevant deontological or prudential considerations, nor any other relevant factors than those implied by the description of the problem.<sup>21</sup>

Problem 1: The first half of the world's history has been like the first half of worlds 2 and 4. At  $t = 1$ , you are offered a choice between bringing about the second half of world 2 and bringing about the second half of world 4. Recall that according to the view now under consideration,  $V2b = V2a = V4a$ . But obviously  $V4b > V4a$ ; hence,

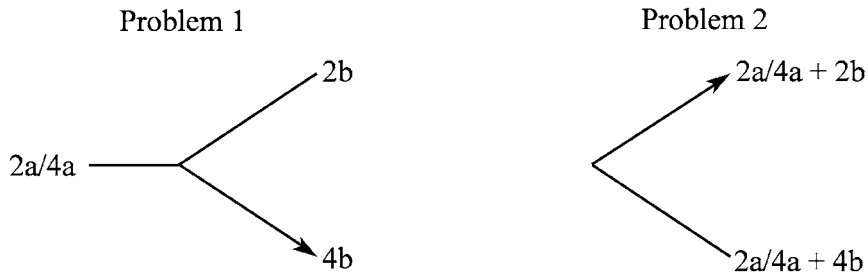


Figure 3. These look like the same choice problem. But denial of Cross-Temporal Additivity leads to the view that you should take  $4b$  in problem 1, but  $2b$  in problem 2.

$V4b > V2b$ . Therefore, you should choose to bring about the second half of world 4.

Problem 2: You are given a choice between bringing about world 2 and bringing about world 4. According to the view now under consideration, world 2 as a whole is better than world 4. Therefore, you should choose to bring about world 2 (see Figure 3).

The answers to these two problems do not seem to cohere. A choice between the second half of world 2 and the second half of world 4, *given* that the first half of each occurs, is equivalent to a choice between worlds 2 and 4. Problems 1 and 2 are really the same problem. But according to the above reasoning, opposite choices are justified in the two cases.

One might claim that there is a difference in that in problem 1, you are making the choice at  $t = 1$ , whereas in problem 2, you are presumably making the choice at or before  $t = 0$ . So modify problem 1 as follows: some time before  $t = 0$ , you are informed that the first half of the world's history *will* match that of worlds 2 and 4, and you are asked to choose only how the second half of the world's history shall be. It seems absurd to suppose that in this case you should give a different answer from the answer you should give if you wait until  $t = 1$  before making the decision.

We turn to a second apparent paradox (see Figure 4). In the first half of world 5, A and B each have horrible lives, with a level of well-being of  $-100$ . Now consider the following choice problems.

Problem 3: Just before  $t = 0$ , you are given two choices in succession. Your first decision will determine (causally) whether the first

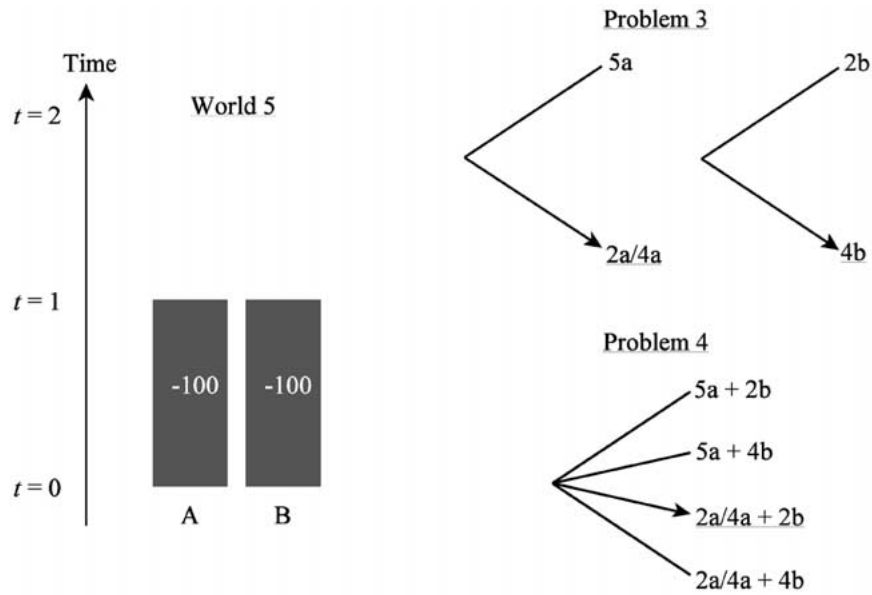


Figure 4. In world 5, A and B each have *negative* utilities. Choice problems 3 and 4 seem to be the same problem, yet the denial of cross-temporal additivity leads to the choice of 4b in problem 3 and 2b in problem 4.

half of the world will be like that of world 5 or like that of worlds 2 and 4. Your second decision will determine whether the second half of the world will be like that of world 2 or like that of world 4. You decide, first, to bring about the first half of world 4, since it is better than the first half of world 5; and second, to bring about the second half of world 4, since it is better than the second half of world 2.

Problem 4: Just before  $t = 0$ , you are given a single choice among four options: The first option is the first half of world 5 + the second half of world 2. The second is the first half of world 5 + the second half of world 4. Third, the first half of world 2 + the second half of world 2. And fourth, the first half of world 4 + the second half of world 4. You choose to bring about the first half of world 2 + the second half of world 2, because this is the best of the four options.

Problems 3 and 4 have the same four possible overall outcomes: the possible joint outcomes of the two choices in problem 3 are the same as the possible outcomes of the single choice in problem 4. It seems absurd that these two problems should yield different results – that it makes a difference whether the two choices of problem 3

are combined into a single choice. Yet, according to the above reasoning, you wind up with world 4 in problem 3 and world 2 in problem 4.

The difficulty can be stated in more general terms. The denial of Cross-Temporal Additivity creates a problem for our intuitive understanding of consequentialist decision-making. *Prima facie*, all three of the following decision rules seem rational (at least for situations calling for impartial consequentialist decision-making in the absence of uncertainty):

- (i) Choose the action which is such that, if you choose it, the world will be best.
- (ii) Choose the action which has the best overall consequences.<sup>22</sup>
- (iii) Choose the action such that, if you choose it, the future will be best.

These rules correspond to three conceptions of what one should see oneself as ultimately choosing among: possible worlds, possible sets of consequences, or possible futures. Normally, barring violations of value additivity, all three conceptions are equally acceptable. But matters are otherwise if we adopt a temporally non-additive theory of value.<sup>23</sup>

Suppose we accept rule (ii), wishing to optimize total consequences. This appears to commit us to the reasoning given above concerning choice problems 3 and 4. The reasoning in problem 4, leading to the choice of world 2, is straightforward. In problem 3, having rejected world 5 for obvious reasons, your second choice has, as possible total consequences, either the second half of world 2 or the second half of world 4. Rule (ii) requires that you prefer the latter. Nor should we argue instead that the choice has, as possible “consequences”, that the world as a whole be world 2 or that the world as a whole be world 4. Intuitively, the consequences of an action cannot lie (even partly) in its past.<sup>24</sup> In particular, if A is some event that had already occurred, and my action causes B, then we should not call the combined event (A + B) an additional “consequence” of my action, over and above B. For example, I can bring it about that I eat a cookie on November 7, 2002. If I do, I will not thereby bring it about that (Napoleon lost the Battle of Waterloo in 1815, and I ate a cookie on November 7, 2002); that fact will not

be a further consequence of my action, over and above the cookie-eating, that I need to take into account in making my decision. Now, since the existence of world 2 just consists in the existence of its first half conjoined with the existence of its second half, and since problem 3 could be modified (without changing anything that seems relevant to the choice problem) so that the second choice is made after the first half of world 2 passes, we should not view the existence of world 2 as a whole as a potential further consequence of this choice in addition to the existence of just the second half of world 2.

Next, suppose we accept rule (iii), wishing to optimize the future. This appears to commit us to the implausible view that, in problem 1, it matters when you make the choice. At  $t = 0$ , since the whole world is in the future, you should prefer world 2. But if offered the otherwise identical choice at  $t = 1$ , since only the second half of the world is then in the future, you should prefer (the second half of) world 4. We could even imagine a situation in which, at  $t = 0$ , you are rational to attempt to commit yourself to later choosing the second half of world 2, knowing in advance that, at  $t = 1$ , it will then be rational to change your mind, even though no change in your information will have occurred.

Suppose, lastly, that we accept rule (i) and reject both (ii) and (iii). In this case, the preceding problems do not arise: problems 1 and 2 yield the same outcome, problems 3 and 4 yield the same outcome, and the time of the choice in problem 1 does not affect the outcome. What, then, is the objection to such a position? Simply that rules (i), (ii), and (iii) (along with the principle of Temporal Additivity) all seem highly plausible intuitively. We can maintain all of them without paradox, provided we renounce Egalitarianism. If it were only rule (i) that seemed initially plausible, then the Egalitarian would escape my paradoxes. But this is not so. Apropos of this, it is worth observing that consequentialist theories are standardly formulated in terms of rule (ii). And to see the intuitive plausibility of rule (iii), consider one more choice problem.

Problem 5: You are free to perform or refrain from performing some action, A. You are reliably informed that, should you perform A, the rest of your life (your future) will be worse than it would otherwise be. However, because of a failure of Temporal Additivity, your life as a whole will have been *better* than it would

otherwise have been. Assume there are no other practically relevant considerations. Should you perform A?

I think there is strong intuitive force to the negative answer. This shows that rule (iii) is at least as plausible as rule (i). I think there is also a sense of paradox to problem 5, showing the counter-intuitiveness of (temporally) non-additive value. Notice that the choice in problem 5 is analogous to the second choice in problem 3 (assuming the latter choice to be made at  $t = 1$ ), with only the difference that problem 5 calls for prudential reasoning, rather than purely moral reasoning. And that difference does not seem relevant to the comparison of rules (i) and (iii).

#### 4. CONCLUSION

The ideal of equality has exerted a powerful and pervasive influence on contemporary moral and political thought. Indeed, Sen claims that *every* major contemporary moral theory is egalitarian in some sense.<sup>25</sup> It is therefore important to assess how much damage my argument does to the ideal of equality. This question naturally divides into two components: First, how many contemporary ethical positions regarding equality does my argument cut against? Second, how should we assess the force of the argument in comparison with the strength of those ethical positions?

It should be evident that there are many forms of egalitarianism that my argument fails to address. It has nothing to say about, for example, the instrumental value of equality of wealth, income, or opportunities.<sup>26</sup> Nor, obviously, does my argument cut against notions of equal rights or equal consideration of interests.

Nevertheless, the anti-egalitarian conclusion I have drawn is a strong one. The thesis of Egalitarianism I have confronted is not committed to holding, for example, that equality of welfare is always desirable *all things considered*, nor does it hold that we are obligated to equalize welfare, or even to do anything close to that. It need hold only that equality of welfare contributes non-instrumentally at least to *some* degree to the value of the world (or that inequality detracts from it), *at least* in certain simple cases, such as that of my worlds 1 and 3.<sup>27</sup>



Furthermore, it seems likely that my argument or one like it could be employed to cast doubt on other prominent forms of egalitarianism. I cannot discuss the matter in detail here, but I would like to at least sketch how such extensions would go. Equality of resources could be shown to lack intrinsic value<sup>28</sup> by means of a thought experiment in which worlds 1, 2, and 3 are as described above and in which (either because the principle of diminishing marginal utility of resources does not hold in this world, or because marginal utility only starts to decline above a resource level that neither A nor B has reached) A's and B's levels of welfare at each time are simply proportional to their resource endowments at each time. It seems that my above argument would still go through, showing worlds 1 and 3 to be equally desirable. One may similarly confront the view that equality of *opportunity* for welfare has intrinsic value<sup>29</sup> by imagining that, as it happens, A and B each capitalize equally well on their opportunities, so that A's and B's levels of welfare in worlds 1–3 are also strictly proportional to their opportunities for welfare.

The argument may even cast doubt on deontological forms of egalitarianism. Thus, consider the Rawlsian alleged obligation to improve the lot of those worst off, even where this does not increase total utility.<sup>30</sup> Given that this obligation is understood deontologically,<sup>31</sup> its existence is not ruled out by the fact that the bringing about of such improvements does not make the world overall *better*. Nevertheless, I think that the latter fact, if it is a fact, must cast some *doubt* on Rawls' conception of justice, because if the Rawlsian conception of justice is correct, and my argument in this paper is correct, then the achievement of distributive justice is not even a good. To see this, imagine that a state had the opportunity to effect a transition from a situation like world 3 to one like world 1.<sup>32</sup> Under Rawls' conception of justice, the state would be obligated to make this transition, at which point distributive justice would be achieved. But according to the argument of this paper, the world would not have gotten better overall, even though there would have been no loss in any *other* value. Therefore, the achievement of distributive justice had no value. If this conclusion seems implausible, we have reason for rejecting the conception according to which the transition was an improvement with respect to justice.

Be that as it may, the argument of this paper certainly contravenes a widespread belief.<sup>33</sup> Is the argument strong enough to overturn that belief, or should we instead react by simply rejecting one or more of the argument's premises, G. E. Moore-style?<sup>34</sup> I suspect that some advocates of equality will be tempted towards the latter response. But to answer this fairly, we need to first remind ourselves of the sort of grounds on which Axiological Welfare Egalitarianism is supported. Arguments for it are hard to come by, and the major argument proponents bring forward is fairly thin: that undeserved differences in utility are unfair and that this unfairness is bad, is typically put forward simply as an unargued intuition.<sup>35</sup> The existence of such intuitions is a reason to support Egalitarianism, but only a *prima facie* reason.

Non-Egalitarianism with respect to utility is not without its own intuitive support: the simplest and most natural account of utility-value is that it just is (proportional to) total utility. It seems intuitive to many (particularly utilitarians) that the value arising from human welfare, satisfaction, or pleasure should add across persons. Perhaps one can make a case that Egalitarian intuitions are more widespread or stronger and therefore at least somewhat to be preferred, but Kai Nielsen's dismissal of the intuitions of utilitarians and others who endorse additivity as "incredible and morally monstrous" is surely out of place. As far as simple appeals to intuition go, I think the case is not at all clear.

Additionally, we have reason to discount (but not entirely ignore) unargued *political* intuitions of the sort that Nielsen appeals to. The reason is that reliance on such intuitions tends to lead different people to widely divergent beliefs, even more so than reliance on intuitions of other sorts, and such intuitions can therefore be concluded to be unreliable and easily biased.<sup>36</sup> Conservatives and liberals often experience divergent intuitions, in particular, about the importance of equality.<sup>37</sup> In contrast, the intuitions relied on in this paper – such as Intrapersonal Non-Egalitarianism, Strong Supervenience, and the correctness of rules (i)–(iii) in section 3.3 – concern abstract moral principles that are not on their face political. For example, one would not initially anticipate that either liberal or conservative attitudes would bias whether one finds decision rules (i)–(iii) acceptable or not. These are the sort of principles that

would be used to resolve disputes in political philosophy, if such disputes can ever be resolved. Therefore, the premises of my argument should be given priority over the fairness intuition and, in the absence of independent grounds for rejecting them, should be taken as refuting Egalitarianism.<sup>38</sup>

## NOTES

<sup>1</sup> Some versions of egalitarianism hold, not that equality is a value, but that people in fact *are* equal in some respect. (Consider the Declaration of Independence's assertion that "all men are created equal".) But these strike me as relatively uninteresting versions of 'egalitarianism', *unless* they are tied to some normative theses along the lines I have indicated (for example, the equality of people in some factual respects might be thought to justify equal treatment of them, to establish the desirability of equality of opportunity, etc.) – at any rate, I shall not be concerned with such versions herein.

<sup>2</sup> I owe the terminology to Sarah Stroud.

<sup>3</sup> Given two things, A and B, the 'fusion' (or mereological sum) of A and B is the thing that has A and B as parts and nothing else (that is, has no parts that don't overlap with A or B). See Leonard and Goodman (1940, p. 47).

<sup>4</sup> Parfit (1991, pp. 3–4) calls this "Telic Egalitarianism". Hereinafter, I shall frequently drop the modifiers, so that "Egalitarianism" with a capital "E" should be understood to refer to Axiological Welfare Egalitarianism, while "egalitarianism" refers to the generic doctrine mentioned at the beginning of section 1.

See Temkin (1986, p. 100) and Nielsen (1985, p. 283), for similar formulations of the idea. There are complex questions about how one should measure the degree of inequality in a situation (see Temkin, 1993). Fortunately, we can bypass these questions here (with one exception, concerning the "complete lives" and "simultaneous segments" views, to be discussed below), since the consequence mentioned in the following sentence in the text holds regardless of the measure of inequality one adopts, and since virtually any egalitarian will agree on the ordering in terms of inequality of the hypothetical situations I will be discussing.

<sup>5</sup> Kagan (1988) offers grounds for questioning this putative entailment, which seems to depend on the suppressed premise (his (p. 12) "ubiquity thesis") that if some factor ever contributes non-instrumentally to the goodness of a situation, then it does so in *every* case in which it is present. Against this, one might hold that the contribution that equality makes to utility-value is contextual, depending on what other factors are present. However, even such a view does not pose a serious problem for my following argument, since the cases I shall consider below (worlds 1 and 3) seem to be *paradigms* of the kind of cases in which egalitarians think that equality/inequality matters; hence, if it can be shown that equality does not matter in *these* cases, it will be reasonable to conclude that equality never matters.

<sup>6</sup> Following McLaughlin's (1995, p. 24) definition of supervenience; Kim (1993, pp. 81–82) confirms that this is equivalent to Kim's own notion of strong supervenience.

<sup>7</sup> In Kim 1976, an event is individuated by a constitutive object, a constitutive property, and a time or time interval of occurrence. Note that my formulation of the strong supervenience principle does not require the assumption that Kim's view is a correct account of *events*, but only that there are some such things as the things Kim calls "events".

<sup>8</sup> The first sentence of this paragraph might be thought, by itself, to support premise 1; this would require invoking (part of) the Person-Affecting Principle (discussed in Temkin, 1987, p. 166; also dubbed "the Slogan" in Temkin, 1993, pp. 248, 256–257): the idea that one situation cannot be worse than another in any respect unless there is someone *for* whom it is worse (in some respect). However, Temkin has presented strong criticisms of this principle, and my argument does not rest on it. Rather, I am arguing that  $V1 = V2$  because world 1 is as good as world 2 for every person involved *and* both worlds are completely egalitarian (and by stipulation all other morally relevant factors are controlled; i.e., we are considering only utility-value).

<sup>9</sup> Almost all thinkers regard time preference, though a real tendency in human nature, as irrational or otherwise bad (e.g., Hume, 1992, III.II.vii; Rawls, 1971, pp. 293–298; Sidgwick, 1981, pp. 124n, 320). An exception is Bentham, who, surprisingly, not only recognizes the tendency but endorses it (1996, pp. 38–41, 169–170), though he does not say why.

<sup>10</sup> An alternative view, perhaps what Bentham had in mind, would be that the goodness of a particular event is time-relative, such that a good event gets *better and better* as we get closer to the time at which it will happen. But this view also seems unmotivated, and it is difficult to understand what is meant by the degree of goodness of an event at times other than the time at which the event exists, where this is different from the goodness the event has when it happens.

<sup>11</sup> Compare the suggestion that  $V3$  is less than  $V1$  because in world 3, B would feel indignation, envy, or frustration about the very fact that his utility is unequal to A's. By hypothesis, if such negative emotions are relevant to B's well-being, then they either do not occur or are already included in B's stipulated level of well-being.

<sup>12</sup> I assume that it is prudentially rational to do whatever improves the overall utility-value of one's own life.

<sup>13</sup> McKerlie (1989) first introduced these notions, followed by Temkin's (1993, chapter 8) discussion. McKerlie and Temkin also discuss a third view, the "corresponding segments view". However, we need not discuss the corresponding segments view separately, since it is equivalent to the simultaneous segments view for the cases I am considering, where the individuals to be compared have equal lifespans and are born at the same time. My objections to the simultaneous segments view apply equally to the corresponding segments view.

<sup>14</sup> McKerlie, 1989, pp. 478–479; Temkin, 1993, p. 236. I confess to finding some intuitive force behind Temkin's claim about his specific case (in which God treats

two equally faithful followers extremely unequally for half of their lives, and then reverses his treatment of the two for the remainder of their lives), but I see no *prima facie* intuitive force to the claim that world 2 is unequalitarian. This may be due to one or more extraneous elements in Temkin's scenario and his presentation, which are not present in my description of world 2. Briefly, these extraneous elements include: a) that Temkin *begins* the description of the scenario with what appears to be an extreme inequality, and only afterwards brings out the information that this inequality is later to be compensated for by an inequality in the opposite direction (unfortunately, the order in which information is presented sometimes affects intuitions – but I do not think any bias could be found in the way information is presented in my description of world 2); b) that Temkin's scenario involves *suffering*, whereas mine only involves different levels of (positive) well-being; c) that God's behavior in Temkin's scenario is clearly unjust independently of considerations of equality – God's harsh treatment of Job2 would be unjust even if Job1 did not exist; d) Temkin's colorful use of the Devil, who gives various evaluatively slanted descriptions of the situation. Contributing to factor (c) (which I suspect is the most important factor) is the fact that, as far as the reader can tell, God's choice to make anyone badly off – or even to not make everyone maximally well-off – seems to be entirely capricious. More reliable intuitions might be obtained from considering a scenario in which a being has only a fixed amount of total utility to distribute – say, 300 points – and he chooses to distribute it as in world 2. It does not seem to me, intuitively, that either beneficiary would have a fairness complaint in this case.

<sup>15</sup> Temkin, 1986, p. 101; Nielsen, 1985, pp. 7–8. Cf. Temkin's (1986, pp. 102–103) remark that whenever inequality exists, there is someone who "has a complaint". Who could be said to have an equality-based complaint about the way world 2 (as a whole) goes?

<sup>16</sup> Temkin (1993, p. 246) stresses this point.

<sup>17</sup> The example is from McKerlie, 1989, p. 483.

<sup>18</sup> Relatedly, McKerlie (1989, p. 488) observes that simultaneous segments Egalitarianism regrets putative inequalities that are not bad *for* anyone; it can lead to the judgement that one history is better than another, all things considered, even though they contain the same people and every such person is better off in the second history than in the first. But see Temkin's criticisms of "the Slogan", cited above in note 8.

<sup>19</sup> Following Kim's (1993, p. 68) definition of "global supervenience". Kim (1993, pp. 82–83) confirms that global supervenience is logically weaker than strong supervenience (as the present example also illustrates).

<sup>20</sup> What if, in response, we say that *both* halves of world 2 gain some 'extra' utility value (relative to world 3) as a result of the fact that they compensate for each other? My answer would be that this position generates the sort of decision theoretic puzzles discussed in section 3.3. Briefly, the reason is that this view allows for the possibility of one's retroactively improving (or otherwise changing the value of) a past event, and any such view will place decision rules (i), (ii), and (iii) below into conflict in certain conditions.

<sup>21</sup> According to some forms of egalitarianism, the stated assumption is impossible, because there is necessarily a deontological obligation. For example, in Rawls' (1971) view, there would be a deontological obligation to prefer world 2. Such deontological forms of egalitarianism are not considered here; for purposes of this paper, I address only the idea of equality as an intrinsically desirable goal. However, one need not embrace consequentialism in general in order to accept my arguments in this section; it is sufficient if one accepts that there is a coherent notion of what it is to decide in a consequentialist manner, and one has certain intuitions about what such decision-making would entail.

<sup>22</sup> Here one may stipulate that "consequences" be taken broadly, e.g., so as to include things one merely allows to happen as well as things one makes happen. We can also treat an action as, in a degenerate sense, a 'consequence' of itself. In each of these rules, "best" means "has the most intrinsic value".

<sup>23</sup> An interesting question here is whether rule (ii) rules out *all* forms of nonadditive value, temporal or not, because it calls for summing the values of an action's consequences. The answer is that rule (ii) need not be read this way; rule (ii) may instead be read as requiring the maximization of the value of (the single event that is) the fusion of all future consequences.

<sup>24</sup> Another way to put the point is to say that what one brings about cannot contain any hard facts about the past. See Fischer (1989, pp. 5–6) on the distinction between hard and soft facts about the past.

<sup>25</sup> Sen, 1992, p. 3.

<sup>26</sup> See, e.g., Dalton (1920, p. 349) for the classic utilitarian argument on this score, and Nielsen (1985, chapter 3), for further arguments for the instrumental value of equality of resources.

<sup>27</sup> The last qualifier weakens the view relative to the definition given in section 1. My point here is that my argument can succeed even against this very weak form of Egalitarianism.

Note that this minimal form of welfare Egalitarianism can escape Dworkin's counter-examples to (stronger forms of) welfare egalitarianism (see Dworkin, 1981a, especially pp. 198, 201–202, 229): stipulate that in worlds 1 and 3, no one's utility or disutility results from such factors as racial prejudice, incorrect beliefs or values, expensive tastes, culpable action or inaction, and so on – or simply that, in general, persons A and B have the same beliefs, desires, and values. In such a case, Dworkin's counter-examples do not obtain; and the *minimal* form of Egalitarianism holds that at least in these, highly restricted circumstances, equality of welfare has at least some value. Even this, I claim, is not true.

<sup>28</sup> Ostensibly contradicting Dworkin (1981b) – though he does not use the language of intrinsic value but instead concerns himself with the question of what counts as "treating people as equals" (1981a, p. 186; 1981b, p. 283).

<sup>29</sup> See Arneson's (1989) view. Cohen (1989) puts forward a similar but more sophisticated view, which appears to be subject to the same objection.

<sup>30</sup> See, e.g., Rawls 1971, p. 62. Note that, though Rawls speaks of inequality of wealth and other resources, the difference principle ultimately calls for a maximization of the *advantage* of the worst-off group.

<sup>31</sup> Rawls, 1971, p. 30.

<sup>32</sup> To spell this out a little more, imagine there are two groups of people in the society, that members of the first group regularly have a welfare level of 100 throughout their lives, and members of the second group regularly experience a welfare level of 50 throughout their lives. The state institutes a redistributive policy which brings everyone's constant welfare level to 75 henceforth. Assume that the state had no Pareto superior policy available.

<sup>33</sup> In an informal survey at the original presentation of this paper, 13 philosophers expressed belief in Axiological Welfare Egalitarianism, 6 expressed disbelief, and 16 withheld judgement.

<sup>34</sup> I allude here to Moore's response to Hume's argument for external-world skepticism: "I *do* know that this pencil exists; but I could not know this, if Hume's principles were true; *therefore*, Hume's principles, one or both of them, are false" (Moore, 1953, p. 120; emphasis Moore's).

<sup>35</sup> See Nielsen 1985, pp. 7–8, appealing to his "root pre-analytical (pre-theoretical) conception of a central element in a good society", going on to opine that it "seems both incredible and morally monstrous" that someone might deny that inequality is unfair.

Sen also gives a very abstract reason, different from the fairness argument, for accepting the value of equality in some sense or other (1992, pp. 16–19); however, Sen's argument does not challenge my position here, as Sen's argument does not support the specific form of egalitarianism I have criticized, and I do endorse other forms of egalitarianism (Strong Supervenience is an egalitarian principle, in Sen's sense).

In addition, Rawls' (1971) argument from the Original Position thought experiment presents a challenge for my position, if the argument of the preceding paragraph in the text is correct, and I cannot detail independent responses to Rawls' argument here. See Nozick (1974, pp. 183–204) and Harsanyi (1975) for some important and persuasive criticisms of Rawls.

<sup>36</sup> There are several strong reasons for regarding political intuitions and other political beliefs as in general easily biased (see Kornblith, 1999; Caplan, 2001, 2002). The most common alleged source of bias in regard to beliefs about equality in particular is that of envy; see Schoeck (1966, p. 231): "[T]he sense of equity, of justice and injustice[ is] inherent in man because of his capacity to envy"; and Hayek (1960, p. 93): "[M]ost of the strictly egalitarian demands are based on nothing better than envy".

<sup>37</sup> Contrast Nielsen's remarks (note 35 above) with those of Hayek and Schoeck (note 36). Though the latter do not purport to express *intuitions*, it is reasonable to infer that neither Hayek nor Schoeck would share Nielsen's intuitions.

<sup>38</sup> I would like to thank Stuart Rachels, David Boonin, an anonymous referee for *Philosophical Studies*, and the participants of the 2002 Bellingham Summer Philosophy Conference, including especially Ben Bradley and Liz Harman, for some stimulating and helpful discussion of the issues raised in this paper.

## REFERENCES

- Arneson, R. (1989): 'Equality and Equal Opportunity for Welfare', *Philosophical Studies* 56, 77–93.
- Benn, S. (1967): 'Egalitarianism and the Equal Consideration of Interests', *Nomos IX: Equality* (pp. 61–78), New York: Atherton Press.
- Caplan, B. (2001): 'Rational Ignorance vs. Rational Irrationality', *Kyklos* 54, 3–26.
- Caplan, B. (2002): 'Systematically Biased Beliefs about Economics: Robust Evidence of Judgmental Anomalies from the *Survey of Americans and Economists on the Economy*', *The Economic Journal* 112, 433–458.
- Cohen, G.A. (1989): 'On the Currency of Egalitarian Justice', *Ethics* 99, 906–944.
- Dalton, H. (1920): 'The Measurement of the Inequality of Incomes', *The Economic Journal* 30, 348–361.
- Dworkin, R. (1981a): 'What Is Equality? Part 1: Equality of Welfare', *Philosophy and Public Affairs* 10, 185–246.
- Dworkin, R. (1981b): 'What Is Equality? Part 2: Equality of Resources', *Philosophy and Public Affairs* 10, 283–345.
- Fischer, J.M. (1989): 'Introduction: God and Freedom', in J.M. Fischer (ed.), *God, Foreknowledge, and Freedom* (pp. 1–56), Stanford, CA: Stanford University Press.
- Harsanyi, J.C. (1975): 'Can the Maximin Principle Serve as a Basis for Morality?', *American Political Science Review* 69, 594–606.
- Hayek, F. (1960): *The Constitution of Liberty*, Chicago: University of Chicago Press.
- Hume, D. (1992): *A Treatise of Human Nature*, Buffalo, NY: Prometheus.
- Kagan, S. (1988): 'The Additive Fallacy', *Ethics* 99, 5–31.
- Kim, J. (1976): 'Events as Property Exemplifications', in M. Brand and D. Walton (eds.), *Action Theory* (pp. 159–177), Dordrecht: D. Reidel.
- Kim, J. (1993): *Supervenience and Mind*, Cambridge: Cambridge University Press.
- Kornblith, H. (1999): 'Distrusting Reason', *Midwest Studies in Philosophy* 23, 181–196.
- Leonard, H. and Goodman, N. (1940): 'The Calculus of Individuals and Its Uses', *Journal of Symbolic Logic* 2, 45–55.
- McKerlie, D. (1989): 'Equality and Time', *Ethics* 99, 475–491.
- McLaughlin, B. (1995): 'Varieties of Supervenience', in E. Savellos and Ü. Yalçın (eds.), *Supervenience: New Essays* (pp. 16–59), Cambridge: Cambridge University Press.
- Moore, G.E. (1953): 'Hume's Theory Examined', in *Some Main Problems of Philosophy* (pp. 108–126), New York: Humanities Press.
- Nielsen, K. (1985): *Equality and Liberty*, Totowa, NJ: Rowman & Allanheld.
- Nozick, R. (1974): *Anarchy, State, and Utopia*, New York: Basic Books.
- Parfit, D. (1991): 'Equality or Priority?' The Lindley Lecture (Department of Philosophy, University of Kansas, Nov. 21).



- Rawls, J. (1971): *A Theory of Justice*, Cambridge: Harvard University Press.
- Schoeck, H. (1970): *Envy: A Theory of Social Behavior*, New York: Harcourt, Brace & World.
- Sen, A. (1992): *Inequality Reexamined*, Cambridge: Harvard University Press.
- Sidgwick, H. (1981): *The Methods of Ethics*, Indianapolis, Ind.: Hackett.
- Temkin, L. (1986): 'Inequality', *Philosophy and Public Affairs* 15, 99–121.
- Temkin, L. (1987): 'Intransitivity and the Mere Addition Paradox', *Philosophy and Public Affairs* 16, 138–187.
- Temkin, L. (1993): *Inequality*, New York: Oxford University Press.

*University of Colorado*  
*Boulder, CO 80309-0232*  
*USA*

