

# **IS UTILITARIANISM INFORMATIONALLY POOR?**

**Keith Dowding\***

October 2002

Department of Government  
London School of Economics  
and Political Science  
Houghton Street  
London  
WC2A 2AE

Email: [k.m.dowding@lse.ac.uk](mailto:k.m.dowding@lse.ac.uk)

\* I would like to thank Ken Binmore, Martin van Hees and Don Ross for their comments.

# IS UTILITARIANISM INFORMATIONALLY POOR?

## I. INTRODUCTION

Most political philosophers are not utilitarians. They were long ago persuaded that the utilitarian calculus produced too many examples that did not fit with our moral intuitions. John Rawls, of course, was one of the key critics in this regard, but Amartya Sen has been the writer who has most extensively attacked the claims of utilitarianism. Sen's views boil down to the claim that utilitarianism is "informationally poor". That is, important facts get left out of the utilitarian calculus which explains why utilitarianism fails to fit our intuitions. In this paper I suggest that this particular criticism directed at "modern utilitarianism" - defined in terms of utility being understood as "choice-based" and not "experiential" - fail. I do not produce a comprehensive defence of utilitarianism here, merely suggest that many of the standard arguments trotted out against it, most of which derive from the work of Sen, fail.

Sen defines utilitarianism in terms of (i) consequentialism, (ii) welfarism, and (iii) sum ranking.<sup>1</sup> Consequentialism is the thesis that the rightness of an act should be seen in terms of its consequences alone. Welfarism is the claim that moral judgements about states of affairs must be confined to the utilities in those states. The utilities of the states are a function of each individual's utility and welfarism takes no account of other factors.<sup>2</sup> Sum-ranking is the thesis that only individual utilities should be simply aggregated by summing together without paying attention to the distribution of those utilities across the population. In this view, injustice consists in the aggregate loss of utility compared to what could have been had we acted differently. Sen believes that the major problems for utilitarianism consist in (a) its distributional indifference, (b) its neglect of rights and other non-utility information, and (c) the fact that it takes no account of the production of utility. That is, it takes no account of why people have the utility functions they have. The basis of all these

claims is that welfarism (and hence utilitarianism) is informationally poor. Because the only information used in the moral calculus is composed of individual utilities other information about individual rights, freedoms and non-utility concerns are ignored.

His arguments against utilitarianism stem from his experiential or “satisfaction” based view of utility. Modern utilitarians and welfarists, assume a choice-based account of utility. Whilst I do not defend at length the choice-based view of utility here, I do show that Sen’s critique of welfarism and utilitarianism depends entirely upon the experiential and not the choice-based account of utility as it is understood by those whose positions he attacks. I argue that choice-based welfarism cannot be informationally poor since by definition it contains all the information that is relevant for the moral metric. I examine briefly why Sen (and Roemer) think that such a theory is “contentless” and consider the importance of this claim. I will suggest that some of Sen’s arguments depend upon conflating two different accounts of choice-based utility in the utilitarian calculus: actual choice behaviour and imagined choice behaviour. One way of measuring utility is to see what people will choose given the situation in which people find themselves. When we make calculations about the marginal value of apples against oranges, or the value to households of clean streets and good local government services, we measure these things given the situation - the wealth, income, social structure and so on - in which people find themselves. We measure the utility people have given who they are, where they are and so on. These measures we use in explanations of economic, social and political behaviour. Such measures of individual utility can also be used in normative contexts. But given that moral and political philosophy is about the way the world should be, not about how it actually is, such measures are not enough on their own. In moral and political philosophy utilitarians also consider the choices people would make given other circumstances. In one version, for example, John Harsanyi makes use of the veil of ignorance device that John Rawls later made famous.<sup>3</sup> He imagines what distributional principles individuals would make if they

had an equal chance of being anyone in society, given that they would want to maximize the utility of society. Thus we need to keep in mind both that modern utilitarianism uses a choice-based conception, and that whilst it may use evidence of individual utility functions from empirically observed behaviour, the main consideration is what utility we imagine people would get under different structures of society.

It is unfortunate that it has become standard to claim that utilitarianism is a “consequentialist” theory where the justness of an act is defined in terms of its consequences. It would be better to describe utilitarianism as a “teleological theory” where the rightness of an act should be seen in terms of the purposes for which the act was designed. We do not ordinarily blame people for the consequences of acts which they did not think of, though we do blame people if their acts have consequences which they did not intend but which they should have foreseen. The nature of that blame however, is different from when those consequences were intended - the acts were *designed* to have those consequences. The distinction between seeing utilitarianism as a teleological theory as opposed to a consequential one may become important when we realize that we may act in certain ways which we know, on average, will have some bad consequences. We still act thus however, because on average we reckon the good consequences outweigh the bad, and it is those consequences, the ones the act was designed to bring about, which justify the act.<sup>4</sup> Perhaps the difference between the terms “consequentialist” and “teleological” is merely rhetorical - the first provides an easier target to discredit with simple arguments - and both can be used as part of the same justification for utilitarianism. But this provides the utilitarian with a good reason for using the term teleological to describe her philosophy, though the term “consequentialism” is so deeply ingrained in moral and political writings it may be too much to be hoped that it can be left aside for the correct term for what is meant - or at least what is meant by those to whom the term is applied.

We also need to keep in mind that utilitarianism might be both an ethical theory and a political one. As an ethical theory, utilitarianism tells people what morally they should do. Roughly speaking it would say: “when faced with any (moral) situation act so as to increase overall utility.” Or more explicitly it would say “always act on those principles which tend to increase overall utility. Only break those principles in special cases when you calculate that there is an unambiguous increase in overall utility by so acting.” As a political theory utilitarianism is about what the state should do. As a political theory it has sometimes been referred to as “government house utilitarianism”.<sup>5</sup> Here the state should execute and implement laws which (tend to) increase overall liberty. The state may give discretion to its legislature, to its executive or state actors be they judges in courts, or street-level bureaucrats making everyday decisions to break certain principles where they are unambiguously for overall utility gain. Generally speaking however, laws are set up to promote the commonweal. Most moral and political writers tend to conflate two notions - the moral and the political - of the utilitarian calculus. Mostly such conflation does not matter. Obviously political theory is based on ethical theory. But whereas individuals may rather easily choose to break principles (“moral laws”, “conventions” and so on) designed to maximize utility overall if they think breaking them better serves some moral (utilitarian) purpose, state actors are more constrained by the laws and conventions which govern their proper behaviour. State actors are less flexible. We should note however, that in its political form utilitarianism may seem more able to overcome the conflict we normally see between liberty and rights on the one hand and utility (or welfare) on the other. Since one very important way in which utility is increased is by allowing people to choose to live their lives how they want, government house utilitarianism implies giving them the freedom and rights so to do. The government house utilitarianism only implies constraining rights and freedoms in cases where overall utility is unambiguously increased though reducing freedom. Those cases are most likely to occur where conflict results, that is where

allowing freedom to act in some ways stops others from acting freely in others. Where the enjoyment of some overwhelms the enjoyment of others.

Modern welfarism and utilitarianism holds a choice-based account of utility usually based on the von Neumann and Morgenstern method.<sup>6</sup> Here individual utility is simply a representative cardinal measure of what an individual would choose. An individual  $i$ 's utility function is constructed for  $i$ 's preferences over lotteries. The function is cardinal, and this entails that when making a choice over  $\{x, y\}$   $i$  cares not only that he prefers  $x$  over  $y$  (respectively  $y$  over  $x$ ) but by how much.<sup>7</sup> Each *utile*, by the construction, is worth an equal amount to  $i$  as every other *utile*. It is usually claimed that we cannot determine whether each *utile* is judged equally by every individual. In positive contexts this does not matter. In normative ones it might. We might say we have no warrant for thinking that each *utile* is “experientially” equivalent for each person. Equally, however, we have no warrant for thinking they are not. This issue is returned to briefly below. In this paper, I demonstrate that under a choice-based account of utility utilitarianism is not informationally poor in the manner that Sen (and others following him) claim. All the pertinent information is contained in the utilitarian calculation. Indeed it must be so, by definition. That is, any information that anyone wishes to claim is important in a moral calculus must be contained in the welfare function. That is what a choice-based account of utility entails.

## II. DISTRIBUTIONAL CONSIDERATIONS

### *Utility and “Experiential Satisfaction”*

Utilitarians argue that despite its sum-ranking approach utilitarianism is not indifferent to the distribution of utility across a society. The standard or Harsanyi defence says that people are essentially the same hence their utility functions will essentially be the same. Thus the law of diminishing marginal utility (the more you have of something the less you value each extra increment) entails that utilitarianism will lead to some form of egalitarian result.<sup>8</sup> Sen recognizes it is possible to define a

metric on the utility scaling such that the social importance of any distribution at the margin is equal to everyone's marginal utility.<sup>9</sup> Together with the assumption that everyone's utility function is essentially the same we get the egalitarian result. Most famously proved by Harsanyi under two procedures, one by making use of a veil of ignorance such that the utilities reflect choices in an original position where each has an equal probability of being anyone, and secondly under assumptions about extended empathy.<sup>10</sup>

Sen believes this egalitarian result for utilitarianism is problematic. He claims that Harsanyi's theorem is merely a "representation theorem".<sup>11</sup> Sen suggests that it is devoid of descriptive content and so cannot form the basis of a moral theory. It is not absolutely clear in the work of Sen why Harsanyi's theorem is supposedly devoid of descriptive content. One reason might be the standard technical problem of interpersonally comparing utilities. We can determine each person's utility function up to a positive affine transformation, that is, one that shifts the origin and changes the scale of the function. Because different transformations may be applied to each person's utility function, two people whose choice behaviour is identical may "in reality" have different utilities represented by the rival transformations. All available evidence underdetermines the correct transformation for each person. Without a theory-independent fact of the matter we cannot say what either's "true" utility function is in comparison with the other. In other words interpersonal comparisons of utility are impossible or indeterminate.

This objection of course relies upon the supposition that *there is* a correct transformation. Roemer produces an example against Harsanyi's aggregation theorem (or rather Weymark's version of it) which relies precisely upon this supposition.<sup>12</sup> Like Sen, Roemer says that Harsanyi's theorems are not "maximizing a *meaningful* sum of individual utilities - a sum, that is, of some measure of welfare which we have initially posited to be interpersonally comparable and summable." He says "suppose that we have a prior conception of interpersonally comparable satisfaction that ... two individuals derive from lotteries" and "Let us suppose for the

sake of simplicity, that the [non von Neumann Morgenstern] utility functions ... are absolutely measurable and comparable” and the utility gained from some outcome “is a specific quantity of satisfaction in meaningful units we cannot tamper with”.<sup>13</sup> In other words, what we have is some independent conception of “experiential” utility or “satisfaction”. We then construct a unique utility function of this experiential utility and compare it to the utility functions constructed using the von Neumann-Morgenstern method. Roemer argues that it is easy to show that they are not equivalent. It is easy to show this because he produces a numerical example where the utility functions of each of this two individuals are different affine transformations of the Harsanyi/Weymark function. Since he has already *assumed* that these transformations are the *correct* ones, he can say that Harsanyi’s function is meaningless even though it is constituted from von Neumann Morgenstern utility functions constructed from choice behaviour that also fit the “true” one. But where does the “correctness” of the transformations come from? It can only come from the assumption that utility is “satisfaction”, or “experiential” which is something that those using choice-based utility functions deny.<sup>14</sup>

Of course we can run a weaker version against Harsanyi and ask why we use this particular transformation and not some other one. Binmore’s answer is that it is the simplest,<sup>15</sup> but a case might be made for different transformation.<sup>16</sup> This would lead to different calculations for aggregate utility based on the same choice behaviour. Though, for Government House Utilitarianism, even if they lead to different institutions, we could still see which institutions individuals would choose allowing a choice between different transforms of their same choice behaviour. The real problem comes when we use different affine transformations for different people. But why should we do that? What warrant do we have for using different transformations for different people beyond the fact that they *could* be used? Only the assumption that utility is really “experiential” and different transformations *actually* captures the experiential utility of different people. But no evidence of such experiential utility is ever offered. Experiential utility appears as an object in political philosophy without

even an attempt at the kind of justification that philosophers of mind once tried for the similar and largely discredited concept of *qualia*.

The seeming reasonableness of very different affine transformations from the same choice behaviour may derive from simple examples. But we must be wary of intuitions derived from simple examples that do not represent the more complex von Neumann Morgenstern utility functions they are used to critique. For example, it may seem reasonable to assume one's person's preference for strawberries over raspberries is stronger than another person's for raspberries over strawberries despite their identically inverse choice over the two items when offered them. But the reasonableness of such an assumption can only be based on *other* behaviour. For example, their facial expressions when eating fruit, or statements made about how much they enjoy strawberries and raspberries in relationship to other things. However, we must recall that von Neumann Morgenstern utility functions represent all choice in all situations, and in that case it seems less reasonable to claim that someone gets more pleasure out of their life than another, despite identical behaviour. One must also recall that modern utilitarianism is about maximizing utility, not pleasure or happiness, or some other "experience" and so is about maximizing what people would choose, rather some other "experience" that they may get when choosing. So the objection is otiose even if it seems reasonable. (Its supposed reasonableness may give a reason for supposing that the "experience" is more important than the von Neumann Morgenstern utility however.) Much more could be said about this difficult issue, but I leave it aside here.

Sen at least at times does not deny that the choice interpretation of utility is meaningful, but he also claims that the "happiness" or "satisfaction" interpretation of utility is meaningful and "neither can *in general* informationally subsume the other".<sup>17</sup> What we should understand by "in general" is not clear. However, I take it that it means we cannot analyse something in terms of a choice-based interpretation of utility that we can in terms of a "happiness" based interpretation. I think this is false. Though it is true, that some uses of revealed preference analysis do not allow

us to make any claims about happiness, it does not follow that whenever we are able to make claims about someone's happiness we cannot do so within a choice-based account of utility.

Sen says that the question over whether I am happier in state  $x$  or state  $y$  makes sense, but is not the same question as to which one I would choose, everything considered. It is true of course, that which state one chooses everything considered is not necessarily the one that makes one "happiest" in the sense that people usually use that term. But is it true that the choice interpretation cannot "subsume" it? To some extent it depends upon how one is using revealed preference analysis. When studying consumer behaviour through aggregate-data analysis one cannot say much about consumers happiness about the states in which they find themselves after choosing some offerings and rejecting others. (But then discovering how "happy" people are in various different states is a notoriously difficult exercise anyway.<sup>18</sup>) But it is not true that theoretically choice-based utilities cannot subsume "happiness". It can be one consideration taken into account in an individual choice function.

Sen says:

If the outcomes are identified *without* specification of mental states in the respective outcomes, then it is not clear why, say, strong independence would be a requirement of rationality, since our choice may be sensibly influenced by anticipated mental states (over which we may have limited control) that would not figure in the outcome specification. If, on the other hand, outcomes do not include mental states as well, then such axioms as strong independence would be almost impossible to apply in practice, since psychological variations of regret, disappointment, relief, and so on, would make states that are otherwise the same different from each other.<sup>19</sup>

However, all we need in a choice-based interpretation of utility is that I may predict my future behaviour and that is what is taken into account when the strong

independence axiom is used. Take Sen's own example from that article. He asks us to consider his choice of activity on reaching home. There are two alternatives,  $L_1$  - a worthy activity such as refereeing an article; and  $L_2$  some self-indulgence such as watching a movie on the video. Say Sen chooses  $L_1$ . Now consider the same choice of activity when at home, but with the history of Sen taking a big risk when going home with some probability  $(1 - p)$  of being seriously injured. If Sen does get home uninjured he still faces the  $L_1$   $L_2$  choice. However, the latter choice describes the mixed lottery  $L_1^* = (p, L_1; 1 - p, \text{hospitalized})$  &  $L_2^* = (p, L_2; 1 - p, \text{hospitalized})$ . Choosing  $L_2^*$  over  $L_1^*$  is perfectly rational since having escaped the great risk of injury Sen might feel justified in some self-indulgence. Sen suggests this should not break strong independence, but then feels that this means we have to take into account mental states. This is just puzzling. A choice-based account of utility simply notes that after certain types of experiences (one of which is as specified in the story) individuals are liable (with some probability noted over some distribution of events) to choose self-indulgence over worthiness to a greater extent than when the individual concerned has not had those experiences. In other words they are not strongly independent. It is simply false that we cannot note this activity without invoking mental states.

What is not in question is that in explaining individuals' actions we use intentional concepts. But this is an entirely different matter.<sup>20</sup> Revealed preference analysis does not invoke reasons for action as such, they are subsumed in the preference ordering (assumed in decision and game theory; surmised in aggregate-data analysis) given the structure of the choice situation. However, interrogating those preferences may allow us to break them down in beliefs and desires as in standard praxeological analysis.<sup>21</sup>

### *The Diamond Objection*

An early objection to Harsanyi's result comes from Peter Diamond.<sup>22</sup> Using von Neumann Morgenstern utilities requires us to assume that utility is not gained through the process of gambling itself. The Diamond objection to Harsanyi's result does not query this assumption as such, but does query whether we will equally value two different *processes* that lead to the same outcome by attacking the "sure thing" principle. Imagine there are two equally probably outcomes A and B which result from the two possible courses of action  $x$  and  $y$  that the state may take. Table 1 shows the results of individual  $i$ 's utility and individual  $j$ 's utility. With  $i$ 's utility first, and  $j$ 's second.

Table 1

World	Effects of Action	
	A	B
Actions		
x	1,0	1,0
y	1,0	0,1

Either action will lead to an unequal distribution of utility between  $i$  and  $j$ : action  $x$  guarantees that  $i$  will get 1 and  $j$  0. Action  $y$  on the other hand will lead to an equal probability that either  $i$  or  $j$  will gain 1 with the other zero. According to the sure-thing principle society should be indifferent between the two actions  $x$  and  $y$ . Diamond objects however, that action  $y$  is preferable to  $x$  because it is fairer. Why does it seem fairer? There might be two reasons. First is that expected utilities are equal under  $y$  and that might be thought to be fairer. Secondly,  $y$  may seem to be a fairer process of choice. Rather than guaranteeing person  $i$ 's gaining 1, it gives both  $i$  and  $j$  an equal chance of gaining 1. The reason why we might think that  $y$  is a fairer

process of choice, of course, may be due to the fact that expected utilities are equal under  $y$  so perhaps we have only one reason, but Diamond mentions both.

If we think this demonstrates that the sure-thing principle is wrong then it provides a knock down argument against Harsanyi's theorems.<sup>23</sup> However, this example trades on our intuitions and subtly uses an experiential and not a choice-based conception of utility. The intuition which leads us naturally to prefer state action  $y$  to state action  $x$  is based upon the assumption that the utilities in Table 1 are mental constructs, rather than choice-based ones. Again we have to be careful that what, in the example, leads our intuitions is relevant to the principle that is being attacked. Deschamps and Gevers suggest that there is no good reason to prefer action  $y$  to  $x$  unless  $i$  and  $j$  know in advance of the outcome  $(1, 0)$  or  $(0, 1)$  which type of action the state had adopted.<sup>24</sup> If  $i$  and  $j$  know in advance which action will be adopted, there will a period of time, before the utilities we see in Table 1 occur, that both  $i$  and  $j$  will derive utilities based on their expectations of the outcomes. Because of this there is a good reason to prefer  $y$  to  $x$  namely the utility that is likely to be gained through  $j$  believing (rightly) there is a chance that she will get the outcome which brings her 1. This is so even if, in fact she gets 0 which would have been guaranteed under action  $x$ .

Deschamps and Gevers' argument does not seem quite right. After all, even if  $i$  and  $j$  do not know that the action was going to determine that  $i$  would get 1 and  $j$  0 before it took place,  $j$  might still be angry if she discovered *afterwards* that that had been the case. Her utility may be 0 if she believed that she stood an equal chance of getting 1 or 0, but after the event if she discovered she was bound to get 0 that knowledge might cause her to get, say -1, since that *knowledge* is relevant to her utility level.

In what way is it relevant to her utility level? It is tempting merely to say that she would feel angry or stressed that she was never in a position to get 1, when under action  $y$  she would have had 0.5 chance of 1. However, that is to equate her utility with an experiential state of mind, which is not what is meant by choice-based utility functions. What we have to say is that she would not have chosen a lottery with an equal probability of actions  $\{x, y\}$  over a lottery with a sure-thing  $y$ . In other words the potential processes are not equal for her and so she would not have chosen them. Nor indeed are they equal for  $i$ 's who might prefer sure-thing  $y$  over lottery  $\{x, y\}$ .

One might object, as Broome does to Deschamps and Gevers that this distorts the example by making action  $y$  distribute the utilities differently to what the example requires. He says "But this treatment robs Diamond's example of its point. The interest of the example is whether action  $[y]$  is preferable to  $[x]$  *even though* both distribute actual utilities equally unequally".<sup>25</sup> However, rather than robbing the Diamond example of its point it queries the interpretation of the utilities in the Table 1. Our intuitions are challenged and are found not to be up to the job because Diamond's example trades on everyone reading it having the intuition that action  $y$  is fairer than  $x$ . But if that is our intuition then that intuition enters into our utility function. Choice-based interpretations of utility require this to be so. It is reasonable to assume that it will also enter into the choice-based utility functions of  $i$  and  $j$ . Broome assumes that the fundamental utility we are considering in the example is utility from outcomes. He says "Expected utility - the desire for an action - is the desire for a means, whereas utility - the desire for an outcome - is the desire for an end, for something wanted for its own sake".<sup>26</sup> But of course for choice-based utility the fundamental utility is for choice over lotteries, *from which* we can derive choice over final outcomes. But we have to be careful what *final outcome* means. In the

same way that statistically any independent variable in one model can also be a dependent variable in another model, any “final outcome” in one choice function, may be a means to another final outcome in another. So we can have utilities over processes since those processes can be elements of a lottery.<sup>27</sup> But surely a “lottery” is a process, and if we can have utilities over processes, then this argument breaks the assumption crucial to the von Neumann Morgenstern method, that no utility is gained through lotteries. We have to be careful here. Person  $j$  is not indifferent, we are assuming between actions  $x$  and  $y$ , despite “our” judgement that the utilities in each world following either actions  $x$  or  $y$  provide equal utility. This should affect “our” judgement over actions  $x$  and  $y$ . But this does not involve any “pleasure” (or indeed choice-based utility) for person  $j$  over lotteries. Her preference for  $x$  over  $y$  is based entirely upon the expected utilities given the states of world A and B provided by action  $x$  and action  $y$ . She might agree, once it turns out that in fact we inhabit World B, that it would have made no difference whether  $x$  or  $y$  is the action chosen. Indeed, had it been assured that we lived in World B,  $j$  would be indifferent between action  $x$  and action  $y$ ; but the example requires that it is not assured.

The Diamond objection thus fails because it trades on our intuitions, subtly assuming an experiential account of utility, by ignoring the fact that we do have preferences over processes. After all, if “we” (including  $j$ ) prefer action  $x$  to  $y$  then that is a utility that needs to be taken into account in the choice between  $x$  and  $y$ . If all the utilities that are relevant are included in Table 1 then indeed we should be indifferent between actions  $x$  and  $y$ , but they are not. The utilities that are not included are the ones represented by our intuitions about the example.

If the utilities in Table 1 do represent fundamental utilities - what states of the world each person would choose - and moreover, the individuals utilities about the

example are to be part of the utility of the final outcomes (which they must be) then the utilities 1 and 0 must already contain information about the choices  $x$  and  $y$ . This means that for  $j$  her utility is 0 when she has no chance of getting 1 and is 0 when she has a 0.5 chance of 0 or 1. The anger and upset we assumed she might feel when she realizes (*ante* or *post fact*) that action  $y$  was taken rather than  $x$  does not occur. It does not affect her utility. This may be hard to imagine, but that is because it is hard for us to imagine such indifference. She really does not care if the probability of the good result goes to  $i$  as a sure thing rather than mixed between her and  $j$ . Person  $j$  must think the same (otherwise his utilities would not be as they are). In that case our intuitions about what the Diamond example show us do not attack the sure-thing principle because our intuitions are not tuned to such reasoning. That might show that no one in “their right mind” would ever be so indifferent. That might show, that equity is part-and-parcel of our intuitions and thus of our utility functions, and so needs to be built into any utilitarian calculus. However, I take it that that is what Harsanyi is trying to do in his representation theorems when he, on the one hand assumes everyone has the same utility function, and on the other builds in the notion of empathetic preferences, in order to drive his egalitarian utilitarian solution.

In fact Sen notes that Diamond’s argument requires that  $i$  and  $j$ ’s utility is comparable.<sup>28</sup> If we add 1 to  $j$ ’s utility we get Table 2

World Actions	Table 2 Effects of Action	
	A	B
$x$	1,1	1,1
$y$	1,1	0,2

Now it might appear that action  $x$  is the ‘fair’ one rather than action  $y$ , since  $x$  gives equal utility in Worlds A and B, whereas action  $y$  gives a 0.5 probability that  $i$  get 0

and  $j$  gets 2 *utils*. This has come about merely because a change of origin in one of the individual's utility function but leaves the aggregate welfare unchanged. Whereas Harsanyi only requires unit comparability, Diamond requires both unit and origin comparability (or 'full comparability') in order to make his criticism.

### *Everyone's Utility Function is the Same*

Harsanyi assumes everyone's utility function is essentially the same. Sen argues (as do many others) that people are not essentially the same. Tastes differ and we should not reward those with expensive tastes by allowing them more than those with inexpensive tastes. However, the expensive tastes objection does not make any sense against a choice-based conception of utility. Harsanyi assumes that everyone's utility function is essentially the same by assuming in his original position theorem a veil of ignorance and in his aggregation theorem extended empathy. We may interpret this to mean that any differences between people will be factored into each person's utility function when they choose the basic structure of society (the distributional principles). There is no rewarding of expensive tastes here. Each will choose the distributional principles given what they would want, and what they know others would want under any conceivable structure. The expensive tastes objection requires that utility is the experiential and not the choice-based concept of modern welfarism.

Of course people have different needs at different stages in their life-cycle, and some people have different needs to others. Of course this implies that they need different resources to reach the same utility levels. But that "same" level is not the "experience" of a given level of "happiness" or "satisfaction". Equalizing utility means in this case - what level of resources would each person give to, say, a disabled person and what to an able-bodied, if each person might be one or the other? What

would one choose to give to each behind the veil of ignorance, or if one had extended empathy for everyone else?

Of course, the veil of ignorance procedure, and the extended empathy arguments are intended to ensure that such differences in utility functions are taken into account to reach an egalitarian result - or more generally to ensure that distributional considerations enter into a utilitarian calculation. In the one case by insurance - we may turn out disabled - and in the second by ensuring that the needs of the disabled enter into the utility functions of others. (In other words it ensures double counting and provides the justification for assuming utility functions are essentially the same.) I discuss double counting below, here all I claim is that it is false that distributional consideration cannot be taken into account in the utilitarian calculus. Clearly it can, since it can be factored into the utility functions of a population imagined to be constructing a constitution for society. Distributional consideration can be brought into the utilitarian calculus in the same manner as they are for Rawls.

### III. LOGICAL AND INTUITIVE PROBLEMS

Even given these considerations Sen has two major arguments against welfarism. The first, and most general, is his famous theorem on the “impossibility of the paretian liberal”.<sup>29</sup> The result demonstrates that the pareto principle - if everyone prefers outcome  $x$  to outcome  $y$ , then the social judgement should be that outcome  $x$  is preferable to outcome  $y$  - is incompatible with individual rights. The proof depends upon accepting as a necessary condition for holding a right Sen’s condition of minimal liberalism. This says that an individual’s right to some outcome  $x$  (relative to some  $y$ ) entails that he must be decisive over whether  $x$  (relative to  $y$ )

comes about. Many have suggested that this condition is far too strong to be seen as a necessary condition for rights-holding, though in the main these social choice writers do not challenge Sen's theorem or dispute his attack upon welfarism. Indeed it is remarkable that in the enormous literature on Sen's impossibility theorem, little of it is actually directed at defending welfarism, the attack upon which was the purpose of the theorem in the first place. Most writers prefer to defend "liberalism" with most common solution that rights-holders can reach pareto optimality if they choose through contractual arrangements. In a partial defence of welfarism I demonstrate that to the extent that paretian liberalism is impossible paretian welfarism is impossible, and to the extent that paretian welfarism is possible so is paretian liberalism.<sup>30</sup> (This does not deny Sen's theorem, it merely questions its relevance as a critique of welfarism. It suggests any solution the liberal may attain the welfarist may attain also.) The reason is that Sen's informal replies to the "contractual solution" to his theorem do not rely upon his account of rights. Rather they rely upon the incentive-incompatibility of contractual solutions that are not rights but power-based. This is not obvious from Sen's account because he defines rights as a power, that is, in terms of "decisiveness".

The only defence that Sen can offer against this position is whether or not we *should* allow such contracts. There are occasions in the defence of his impossibility theorem that Sen seems to suggest that we should not, though in a reply to Brian Barry says this is not the important issue in his proof.<sup>31</sup> However, Sen's views on the morality of certain types of preferences being counted in the utilitarian calculus (or rather being counted in theory of justice) are clear. They constitute his second major attack upon welfarism (and so utilitarianism).

The second major attack upon welfarism relies upon our intuitions about a simple example and fails for much the same reasons as the Diamond objection to Harsanyi's theorem. Sen's example involves two individual's whose utilities are compared.<sup>32</sup> In Table 3 we are asked to consider the following utilities:

Table 3

	Person A	Person B
World 1	12	5
World 2	10	9
World 3	10	9

We are invited to think that World 2 is preferable to World 1 since redistribution has brought poor person B almost up to the utility level of person A (to simplify imagine these resources are composed entirely of food and in World 1 person B goes hungry, but is satisfied in World 2). World 2 also has greater aggregate utility (through diminishing marginal returns) though here this may not be thought to be important. However, we also have World 3. Here the welfarist is forced to say that this is morally equivalent to World 2 since the utility information is the same. In World 3 there has not been a re-distribution of resources, person B remains equally as hungry as in World 1, but rather is "allowed" to torture person A which brings great pleasure to B, whilst person A's utility is reduced through the pain the torture brings. Welfarism and hence utilitarianism are informationally poor because they cannot distinguish the two cases, since it does not take into account where the utility comes from. Furthermore, some preferences may need to be discounted entirely on the grounds that they are immoral. Whilst "laundering preferences" may be a solution to this problem Sen believes that preference laundering is essentially the same as using non-utility information.<sup>33</sup> In other words our reasons for deciding which preferences to launder will use non-utility information. However, we note that this example

requires a non-choice based account of utility. Only if persons A and B truly cannot judge the difference between sharing food to reach the utility levels of World 2 and person A keeping his excess food and B torturing him, then and only then, are the two worlds identical with regard to the utility information they contain. What this means for choice-based utility is that both A and B, when offered a choice between two equal lotteries differing only in the fact that one lottery contains a world where A has less food and is not tortured, and B has more food but cannot torture A, and in the second A has more food but is tortured, and B is hungry but gets to torture A, then both can only choose between the lotteries by tossing a fair coin. It is simply false that the details of the two worlds are missing from the utility information, since the utility information is the choice made by the individuals given the details of the two worlds.

Sen says, even then we should judge World 2 preferable to World 3. But on what grounds? We are not in either world. If we are not in either world why should our intuitions count? If our judgements are to be brought in, then that also brings in new choice-based utility information. If we prefer World 3 to World 2 then a choice-based welfarist account will do so too. It must do so, because person A and person B are indifferent between Worlds 2 and 3, and if we are not indifferent then our utility will tip the balance. If our preferences are being counted along with that of person A and person B then Table 1 must be re-written to take that into account.

But what if we rewrote Table 1 such that both A and B prefer World 3 where B gets to torture A? In that case, in a utilitarian calculation, “our” preferences for World 3 over 2 may not tip the balance in favour of 2. Does utilitarianism get the wrong answer then? Well it suggests an answer that does not fit with Sen’s intuitions nor probably with few others’ either. However, there is nothing inconsistent in the welfarist argument, which was, after all, designed not to allow the personal preferences of the moralist (or what Bentham called *ipsedixitism*) to hold sway.<sup>34</sup> The whole point of welfarism or utilitarianism is that each individual’s views should be counted equally in the moral calculus. The fact that the moral calculus will

provide answers some people (us) do not like is hardly a criticism. I believe there are good reasons why stable systems of justice will entail the choice of World 2 rather than World 3 - in other words why we have those intuitions about this example and which allows us to make a distinction between mere tastes and moral judgements - but to go into this in any depth will take us too far afield from the critique of the informational paucity of choice-based welfarism and utilitarianism.

We may bring in other considerations into this example. We may worry about Poor being able to carry out his other responsibilities when hungry and so on, we may worry about the future of a society where torture is considered on a par with feeding the hungry, and we may feel it is wrong to develop preferences for torture. But all of these other considerations can also have welfarist representation. In the first we may ask to whom has Poor got responsibilities that he can best carry out if he is well fed? That either brings in other people (and their welfare) or brings on other non-utility information of a different sort, such as the “welfare of the planet” - imagine Poor is a gardener and he can’t water his plants if he is too hungry or too busy torturing Rich.<sup>35</sup> One way of worrying about the future of society is to worry about the welfare of people in this and future societies (the health of Poor’s children).

#### IV. THE PRODUCTION OF UTILITY

At this point we may enter a critique of utilitarianism that concerns Sen’s third criticism. Utilitarianism does not take account of where preferences come from. Sen suggests that the poor accept their poverty getting more utility from their life than the rich would living that life. In terms of the example above we may worry about people gaining a preference for torture. Such a worry seems to be anti-welfarist in the following sense. Utilitarianism only examines the utilities of the individuals and takes no account of other information. Where those preferences come from is surely other information. However, again, in Harsanyi’s original position account we may be worried about the generation of preferences. We may wish to consider what sort of people we want to be. In particular to give a welfarist spin to this discussion we

might consider what sort of people will be good for a future society we may be a part of. Even here, our intuitions enter in and that means utility considerations. In other words, we may consider, how stable a society may be with one set of preferences as opposed to another. Will a society where people prefer to torture others than eat well, be a longer-lasting society than one where people prefer to eat than torture? This looks like the sort of question that may be asked in evolutionary contexts. And so it should. Sen's example trades upon our intuition that torture is bad and ensuring that people eat well is good. These intuitions are not generated simply happenstance. We have them for reasons, and it is right and proper that these reasons should enter into our consideration of the example. Why should we not expect to see a society where World 2 is considered on a par with World 3? The reasons why we should not expect to see such a society are precisely the reasons why we have the intuitions we have about the superiority of World 2 over World 3. We should not use our intuitions about these worlds against welfarism without examining the example more carefully.

I am by no means producing a full-blown defence of utilitarianism here. Rather I am giving a fuller discussion of an example trotted out against utilitarianism by Sen and others which has not received the attention it deserves simply because it seems so obvious that we should prefer a world where people are not tortured to one where they are. But if we do think this we think it because that is our intuition and saying it is our intuition is another way of saying that it enters into our utility function, and if we take account of intuitions when assessing the example we are taking into consideration our utility function(s) and hence cannot claim we are assessing the example on non-welfarist criteria. If we strictly consider the example outside of our intuitions, we consider what sort of world the two individuals would inhabit given their utility functions, the moral worth of our intuitions is irrelevant. Either we may find ourselves discussing the development of preferences that is only arguably non-welfarist. Or we may find ourselves having to justify why certain "non-welfarist" information should be taken into account. If we find that torture is something that we do not want, our justification may end up with welfarist

credentials. Alternatively we may simply add some other criteria that is claimed to be non-welfaristic - the welfare of non-humans for example. Can this be made compatible with welfarism?

Robert Goodin, for example, argues that green values are simply a source of utility that might otherwise be overlooked. His theory may be summarized:

- (1) people want to see some pattern to their lives,
- (2) that requires that their lives be seen in some larger context,
- (3) the products of natural processes untouched by human hands provides that larger context.<sup>36</sup>

Even here then, the world outside of humans, which, many believe, should be factored into accounts of what we should do, is mediated through individual utility considerations. Goodin is aware that not all greens would accept this version of a green theory of value. Brian Barry suggests that anyone who holds an ecocentric conception of the good would be right to reject Goodin's formulation. This is because Goodin has no comeback to anyone who said that after due consideration they would sooner have cheap hamburgers to the preservation of the Brazilian rainforest or that argues that the only reason for banning stag or fox hunting is the utility considerations which go into the calculus from those who enjoy the activity and those who are offended by it and not "the obvious objection based directly on the suffering of the foxes, stags, and hares."<sup>37</sup> I think the issue is more complicated and that moral considerations have a deeper connection to individual utility considerations than Barry allows here.

In what sense could it be wrong to hunt foxes? In terms of moral psychology it can only be wrong when we think it is wrong. Is it wrong to whistle at the wind? In what sense could it be wrong to whistle at the wind? Again, in terms of moral psychology it can only be wrong to whistle at the wind when we think it is wrong. Why should we think it is wrong to whistle at the wind? I do not know. Perhaps because it is a waste of time, the tune we whistle gets blown away doing no one any good, and we should be spending our time doing someone some good. (Of course if I

enjoy whistling at the the wind it does do someone some good, and it might do others good, if I feel better about myself for doing it and so do better for others too.) I do not want to get involved in a silly discussion of a silly example. I just want to make a not-so-obvious point. Reasons must enter into why anything is thought to be wrong. If it is wrong to hunt foxes then there must be reasons for it. I take it that one of the strongest reasons why we should not hunt foxes is that it causes them great suffering. Indeed that is the reason why the evidence of veterinary scientists have been important in Britain to the parliamentary considerations of whether to ban stag and fox hunting, and the differences (if any) between these and fishing. The fact that some people are upset by these activities is important, but is not the only consideration for the utilitarian parliamentarian. Of course, in practical democratic affairs, the numbers of people for and against a practice are of prime importance to whether legislation is likely to be proposed by an executive and passed by a legislature. But that is, to some extent another matter. On the moral issue, when I am considering my stance on fox hunting others views are important but not of prime importance. What matters are the moral reasons as I see them. Do animals suffer and how important is their suffering? I take it there can be many arguments made to persuade me one way or the other. Thus non-welfarist information - the suffering of other animals, the nature of the destruction of the planet - enter into my calculation. But that does not make my calculation non-welfarist. For my decision based on these considerations - and every other non-welfarist information such as the interactions of different chemicals (when considering drug abuse), the force of gravity (when considering blame following brick falling from a gantry at a building site) and so on - goes into every moral judgement I make. My judgement in all cases will form part my choice-based utility function. The utilitarian calculus, certainly the “government house utilitarian calculus” might be based on the summation of everyone in the society who has gone through that process. Will the government house utilitarian calculus get the “right answer”? Not according to me if the government decides not to ban fox hunting and I am opposed to it. Binmore would say, that I have no warrant to claim this, and in

doing so am simply revealing myself as an ipse-dixitist. The moral requirements are the result of the social calculus. If I think the moral requirements of the utilitarian calculus are wrong that simply shows my “taste” differs from the utility function (the “values”) of society. Binmore uses Bentham’s ipse-dixitist argument as follows. Individuals own views about justice are “tastes”. We each have our own views about what just distribution and correct moral behaviour are. These are to be distinguished from “values” which are what society sees as just distribution and correct moral behaviour. Of course society is not a being with tastes or values, and so the values of society are some form of conglomeration of the tastes of the members of society. We could see this conglomeration in two senses. First social values could be seen as the aggregation of all individual tastes chosen through some aggregation mechanism. This is straightforwardly some form of welfarism or utilitarianism. Secondly, and closer to Binmore’s own naturalistic account, the values could be those seen in the equilibrium strategies that individuals in the society adopt in order to lead fruitful lives; less straightforwardly utilitarian but welfaristic nonetheless. The institutions, which constitute these equilibria, and the maxims or moral principles that guide actors into those equilibrium strategies would constitute the values of society.<sup>38</sup> However, Binmore’s complaint about those who wish to criticise the current equilibria, or think that the results of the aggregation mechanisms we have are unjust and wish to construct new maxims and moral principles to try to create new equilibrium strategies makes little sense. Stating that because I argue for new criteria is to merely to propose my own prejudice suggest that my views are not ones that I have developed on reflection. It is not true that all political and moral philosophers who criticize current social values are merely elucidating tastes over which they have not reflected. They have not pre-judged the issues, they have judged them. It is true that their arguments may be based upon their tastes or moral intuitions, and it is true that they wish to change the tastes or intuitions of others. It does not follow that they cannot be discussing moral values however. If I can persuade others that my tastes are superior to theirs and am able to change theirs, then both the aggregation mechanism and the

equilibrium strategies may churn out new values. Thus one can affect values by arguing for tastes.

Binmore states:

Tastes are held by individuals. Values are held in common by society. *I* have tastes - *we* have values. ... How are values confused with tastes? Two sources of error are mentioned here. The first occurs when people try to get special treatment for their personal tastes by calling them values. The second arises when appeals to folk versions of the principle of revealed preference result in values being mistaken for tastes.<sup>39</sup>

But a moral naturalist like Binmore should be able to see that *if* I can get my tastes accepted by others then my (now shared) views determine the aggregation mechanism, or if my tastes become the equilibrium strategy by persuading others to act on them - then they automatically become values. It is true, and part of Binmore's two-volume treatment of social justice exposes this, that some tastes cannot become equilibrium strategies, but that is a separate argument from the charge of ipsedixitism. That is an argument that shows some views (tastes) cannot (successfully) become the moral or political principles of society - ought implies can. The question of whether we could, as a society, prefer torture to eating well, in other words, whether Sen's example can be an example of a welfare function need not be answered. If it can, then "our" preferences (that is, people who are not in World 2 or 3) are irrelevant. If it cannot, then this may help explain why we have the intuitions we do.

The second problem alluded to in the Binmore quotation - the "folk version of the principle of revealed preference" is harder to explain, but the objection looks like the following. Choice-based utility accounts do not claim that utility provides reasons for acting. Utility does not provide reasons for anything. A Utility function is a mathematical device for describing a person's actions. The re-description is useful because it allows us, via several consistency postulates, to make predictions about that person's behaviour that are not witnessed or are never carried out. Binmore describes the basics of revealed preference analysis as covering three sets, A actions,

B possible states of the world, and C final consequences (of A given B). Revealed preference theory provides the consistency conditions for a person's actions (in A) by saying she acts as though maximizing the expected value of a utility function (defined in C), relative to a probability distribution (defined in B). Any model that is used to describe the person's actions has to be chosen carefully such that everything defined in C operates given the distribution defined in B. That means that everything a person takes into account must be modelled, including social norms and expectations, and not simply personal desires and tastes.<sup>40</sup>

It is perfectly reasonable for someone to claim that the utilitarian calculus, or the mores of society (the equilibria in the game of life) have got the wrong answer. This is simply part of the attempt to change the equilibria and change the calculus by persuading others that they have got it wrong. And that is how moral values differ from tastes strictly understood, they only count as moral values if we have reasons for holding them, but tastes proper are things we just have.<sup>41</sup> The philosopher that generates a theory that allows us to see that some outcome is to be socially preferred, and becomes socially preferred, can say that his theory was genuinely a theory of value, and not just an expression of tastes.<sup>42</sup> Many things in that are not "welfarist" enter into an individual's moral calculation. Once that individual calculation has been made, her utility function can be plugged into the social welfare function. Even if the social welfare function does not produce her preferred outcome this does not mean that her preferred outcome is not a moral consideration (just a taste); and she may well still have a warrant for claiming that the social welfare function has got the wrong answer. Her reasons for claiming that will be the same reasons why she reached her own view in the first place, and she has every warrant for trying to persuade others about the rightness of her views - and every warrant to use terms such as rightness, wrongness, 'moral imperative' and so on to help persuade others. If she persuades enough then the utilitarian calculus may well shift to her view - that may constitute a new equilibrium. (Only may, since her view may not be a sustainable position in the long run, it may not be part of an equilibrium.) If the utilitarian

calculus is thus the government house utilitarian calculus may also be thus. But I do not think the two will always be the same, and not simply for the practical reason that governments do not always do what the majority want but are bought off by special interests. The government house utilitarian calculus may differ for *moral reasons* from the strict utilitarian calculus of each persons view. How so?

Governments, even democratic ones, do not always do what they know a majority of the public wants: free votes on capital punishment in the British parliament is a testimony to that. Public opinion polls have demonstrated that the public generally speaking is in favour of capital punishment, but the votes for it are always lost, usually by a much bigger majority than commentators think they are going to be. The justification for a free vote, and for members of parliament (MPs) not taking into account of their constituents views is that this issue is one of moral conscience. There is no legal or conventional obligation for any MP, once elected, to act as a delegate and may vote any way they choose (or not bother to turn up at Westminster at all). This may seem that MPs simply vote their “tastes” in Binmore’s terms. However, they, and governments’ more generally may have moral reasons which still fit government house utilitarianism for doing things which they know that a majority of people oppose. A government may argue that what it does is, in its calculation, the best thing for the people. It may argue that even though most people oppose the policy, the policy is right, and right for them. They may argue that they know better than the people because they are more able to calculate the best interests of the people. The government has fuller and superior information, has considered at length the pros and cons and come to a decision. This kind of argument be used about fox-hunting as equally as one about complex issues such as economic policy. No one thinks it odd if government raises taxes, or does not increase spending on education or health, despite knowing that a majority of people are opposed to tax rises, or want more money spent on education or health. The government may argue that the popular policies will, in the long-run have a deleterious effect on the economy making everyone worse off. This, straightforwardly, is a utilitarian calculation. But in what

sense can fox-hunting be bad for a person who enjoys hunting, particularly in that person is in the majority with regard to the propriety of fox-hunting? A government could argue that whilst most people are in favour of fox hunting, the government knows best because it has collected more evidence about the cruelty involved. It might say, that if everyone (or enough others) had this information then those people would think it immoral. This means it is right to ban fox hunting even when there is a majority in favour of continuing the practice.

It might seem that to claim that such an argument is “utilitarian” is tortuous beyond belief. The argument depends on a government house calculation of what people would believe if they had the information the government has. Since this seems equivalent to doing what the government deems “best” it hardly seems to be utilitarian at all. But this is no more than a practical government house version of Harsanyi’s veil of ignorance theorem where individuals choose having full information.

The point of all of this is not really to argue that any view can be utilitarian, rather it is to suggest that the utilitarian position cannot be so easily dismissed as Sen and others do with their rather simplistic examples. Of course, this defence of utilitarianism has been based upon the understanding of choice-based utility, albeit “choice-based” only in a counterfactual (“imagined”) sense. Sen and Roemer’s dismissal of choice-based utility constituting a meaningful utilitarian position denies any relationship of the argument here as being one which can be defendably “utilitarian”. One way of arguing is to define your terms and defiantly stand your ground. I would argue that the heritage of the utilitarianism I am suggesting is that of Bentham and Mill.<sup>43</sup> Another is to ask what is wrong with the choice-based account. There are two general answers. One is that what goes into choice is everything we consider no matter where our views come from. That is, those whose views are not “utilitarian” get counted in the utilitarian calculation. But that of course is precisely why utilitarianism might be presented as an advance over ipse dixitism. All get

treated equally in the calculation and ipsedixitism is avoided. Another response is that we may find there is “double-counting”.

## V. DOUBLE COUNTING

Individuals may strategically use the measuring of their utility functions for their own advantage. Consider the following example used against the choice-based account of utilitarianism.<sup>44</sup> Imagine a cake to be divided between Desdemona and Iago. Iago, being egoistic would take all the cake for himself, but Desdemona, despite her dislike of Iago, is fairminded and would split the cake equally. Should we split the cake 75:25 as a compromise between how each would choose? Gibbard says not on the grounds that there is “no parity” between Iago’s selfish choice and Desdemona’s commitment to a fair outcome. Sobel says that a utilitarian using choice-based utility cannot be so sensible since Desdemona’s motivation for fairness must include Iago’s own view of his own utility. Must it? It is true that if both Desdemona and Iago were to vote for how the cake is to be distributed in that manner, and then the utility compromise reached would be the 75:25 division, then Desdemona (and of course Iago) may be feel dissatisfied that they did not get what they voted for. But Desdemona may feel more dissatisfied because she may feel cheated, whereas Iago may feel that he has put one over the hapless Desdemona. What matters of course, is whether Desdemona *would have voted* for an even division, if she had known Iago was going to vote it all for himself and the compromise would be a 75:25 division. If she still votes for an even division then we can argue that the choice-based utilitarian compromise is indeed a 75:25 split, but if she were to vote it all for herself *in order to gain* the 50:50 division, then the choice-based utilitarian account would see that as the compromise. Indeed in experimental situations fair-minded people do retaliate in iterated games against those who try to cheat in this sort of manner, and, if the game allows, punish the cheaters.<sup>45</sup> Does this mean that Iago’s utility is counted for less than Desdemona’s because she gets what she “really” wants and he does not? No, for when Desdemona knows how Iago is going to vote then her vote is her “choice-

based” utility. Iago is in the same position. Even when he knows what Desdemona is going to vote for, he will still vote it all for himself. Though it is also true that if Iago knows how Desdemona is going to vote, if she knows how he is going to vote, then Iago will know that if he votes for an equal division and Desdemona knows this, Desdemona will vote for an equal division. In that case he may as well vote for an equal division as well (and get the credit for being fair). With complete information they may both vote it all for themselves, or both vote for an equal division, the result will be the same. On the basis of the outcome of this vote alone they should be indifferent between these two sets of votes. Though they may not be indifferent between the underlying psychology of Iago and Desdemona. It is true, of course, that Iago would still sooner vote it all for himself if Desdemona were to vote for an even split, and that Desdemona prefers an even division, and, in theory, we can know this from their bets over lotteries.

It is true, as Dworkin points out, that preference aggregation where some take into account others’ views when voting, but others only consider only their own views, may involve some individuals views being “counted twice” in relation to others.<sup>46</sup> But that is only a part of a choice-based account of a given utilitarian outcome. When the vote is made even when everyone knows what the outcome will be. That is, when someone still chooses to vote taking into account others’ views even when those others will not take into account hers. Sen does not like this. One of his examples to illustrate this sort of behaviour is the relative health of women in parts of India. In difficult times, women will give up food for their male partners, their children and give priority in food and health distribution to sons over daughters. Sen wants an egalitarian solution, where the needs of men and women are considered equally, no matter what the people in that situation think. Of course, the Harsanyi egalitarian utilitarianism overcomes this problem by assuming that individual utility functions are similar, and so people will not give up food for others by choice. Or rather, it assumes that utilitarian institutions will be set up, and the distributional principles adopted under those institutions will be egalitarian. Once the institutions

are in place, however, people may choose of their own free will to disturb the patterns the egalitarian desires. We might think this is the fundamental conflict between rights and freedoms on the one hand, and welfare and equality on the other. Perhaps this may be a problem for utilitarianism (though I am not convinced it is more of a problem than, for example, Rawls' theory or Sen's capability approach) but even if it is, this is not due to informational paucity. If we have a calculus of choice behaviour when some sacrifice themselves for others in full understanding of what they are doing, then we have some sacrificing themselves for others with full information. They know what they are doing. I suspect that the Indian women who give up food or health care for their husbands and children, or give more to the male children than the female, do so knowing what they are doing. Sen does not approve, nor do I, but I think we need to explain our intuitions here. And that again takes us into naturalist ethics. I do not want to say much about that, but let me just suggest the following.

Often when we find cultural practices strange, it is because we do not understand the reasons behind them. Some of course are simply conventions, different equilibria from those our own society has found for the same or similar collective action problems. But often the problems are (perhaps subtly) different which helps explain the practice. When we realize the problem we may discover the reasons why the practice occurs. They may even be what we would do if we were in that position. Often, not only for alien cultural practices but also for our own, the reasons for the practice no longer hold. Their purpose, or function, has disappeared since the environment has changed. It may well have been rational for Indian women to give up their health and that of their female children for those of the men when the health of household, including the women, depended more upon the men than the women. These days, not only in our culture but the cultures Sen is considering that is no longer the case.<sup>47</sup> We can understand our intuitions, their intuitions, and feel that ours are now superior. We can justify our views theoretically. For that reason, to *argue* for a specific value-judgement in a given case is not simply ipse dixitism. We may be trying to show that the values generally held in society are not the best

equilibrium strategies, or even that they are some sort of cultural holdover from the past which are not strictly in our interests. The only *reason* why some set of values might be preferable is the reason provided by arguments. For a naturalist like Binmore, the selection criteria may be based on Dawkins' meme replication,<sup>48</sup> but even then arguments may still be important in that replication process. My argument in this paper, is that we can do so with the tools of utilitarianism just as well as with other moral approaches. Or more correctly, I am not convinced we cannot, and I am particularly not convinced we cannot because choice-based utilitarianism is informationally poor. The reason for my belief is simple. Any information that goes into any moral theory, can also enter into a choice-based utilitarian one. Indeed, it must do.

---

<sup>1</sup> It would be too much to cite all of Sen's works in this regard, but two key articles are "Utilitarianism and Welfarism" *Journal of Philosophy*, LXXVI, 9 (1979): 463-489 and "Informational Analysis of Moral Principles" in *Rational Action: Studies in Philosophy and Social Sciences*, ed. R. Harrison (Cambridge: Cambridge University Press, 1979): 115-132. See also his recent *Development as Freedom*, New York: Oxford University Press, 1999.

<sup>2</sup> We could give a more technical definition of welfarism where a social choice function is welfarist if it satisfies weak ordering, independence and Pareto indifference. In the context of the arguments of this paper, nothing much hangs on the technicalities.

<sup>3</sup> John C. Harsanyi, "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking" *Journal of Political Economy*, LXI (1953): 434-435.

<sup>4</sup> I am not making the distinction between "act" and "rule" utilitarianism here, unless "act utilitarianism" is something nobody has ever believed in, since even act utilitarians are teleological in this sense.

<sup>5</sup> Robert E. Goodin, "Government House Utilitarianism" in ed. Lincoln Allison *The Utilitarian Response: The Contemporary Viability of Utilitarian Political Philosophy*, (London: Sage, 1990): 140-160; Goodin, *Utilitarianism as Public Philosophy*, (Cambridge: Cambridge University Press, 1995) or "institutional utilitarianism" by Russell Hardin, *Morality Within the Limits of Reason*, (Chicago: University of Chicago Press, 1995).

---

<sup>6</sup> John von Neumann and Oskar Morgenstern, *The Theory of Games and Economic Behavior*, (Princeton: Princeton University Press 1944); John Harsanyi, “Normative Validity and the Meaning of Von Neumann and Morgenstern Utilities” in *Studies in the Logic and the Foundations of Game Theory*, ed. Brian Skyrms (Dordrecht: Kluwer, 1992): 266-278; Ken Binmore *Fun and Games: A Text on Game Theory*, (Lexington, Mass.: D. C. Heath and Co, 1992); *Game Theory and the Social Contract I: Playing Fair*, (Cambridge, Mass.: Massachusetts Institute of Technology Press, 1994), *Game Theory and the Social Contract II: Just Playing*, (Cambridge, Mass.: MIT Press, 1998).

<sup>7</sup> Exactly what this means in a choice-based construction is often misunderstood. See Binmore *Just Playing*, pp. 525-26.

<sup>8</sup> John C. Harsanyi, “Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls' Theory” *American Political Science Review* 69 (1975): 594-606.

<sup>9</sup> Amartya Sen, “Equality of What?” in *Choice, Welfare and Measurement* (Oxford: Blackwells, 1982): 353-369, pp. 354-356.

<sup>10</sup> John C. Harsanyi, “Cardinal Utility in Welfare Economics”; “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility” *Journal of Political Economy* 63 (1955): 309-21. For criticisms and amendments see Peter C. Fishburn, “On Harsanyi's Utilitarian Cardinal Welfare Theorem” *Theory and Decision* 17 (1984): 21-28; John A. Weymark, “A Reconsideration of the Harsanyi-Sen Debate on Utilitarianism” in *Interpersonal Comparisons of Utility*, eds Jon Elster and John Roemer (Cambridge: Cambridge University Press, 1991): 255-320; John E. Roemer, *Theories of Distributive Justice*, (Cambridge, Mass.: Harvard University Press, 1996),

---

ch. 4; Peter J. Hammond, “Ex-Ante and Ex-Post Welfare Optimality Under Uncertainty” *Economica* 48 (1981): 235-250, “Ex-Post Optimality as a Dynamically Consistent Objective for Collective Choice Under Uncertainty” in *Social Choice and Welfare*, eds P. K. Pattanaik and M. Salles: 175-206 (Amsterdam: North-Holland, 1983), “On Reconciling Arrow's Theory of Social Choice with Harsanyi's Fundamental Utilitarianism,” in *Arrow and the Foundations of the Theory of Economic Policy*, ed. G. R. Feiwel (London: Macmillan, 1987): 179-22;. Kim C. Border, “More on Harsanyi's Cardinal Welfare Theorem,” *Social Choice and Welfare* 2 (1985): 279-281; John Broome, “The Bolker-Jeffrey Expected Utility Theory and Axiomatic Utilitarianism,” *Review of Economic Studies* 57 (1990): 477-502; Binmore *Just Playing* 1998.

<sup>11</sup> Amartya Sen, “Informational Bases of Alternative Welfare Approaches: Aggregation and Income Distribution,” *Journal of Public Economics* 3 (1974): 387-403; “Non-Linear Social Welfare Functions”, in eds R. Butt and J. Hintikka *Foundational Problems in the Special Sciences*, (Boston: Reidel, 1977): 297-302; Roemer, *Theories of Distributive Justice*.

<sup>12</sup> Roemer *Theories of Justice*; Weymark “An Reconsideration”. This theorem is also called the “Utility sum Theorem” by Amartya Sen “Social Choice Theory” in *Handbook of Mathematical Economics III*, eds Kenneth Arrow and Martin Intiligator (Amsterdam: North-Holland, 1986); “Teleological Utilitarianism” by Binmore, *Just Playing*; *Playing Fair*; and “the Axiomatic Justification” by John C. Harsanyi, “Morality and the Theory of Rational Behaviour” in *Utilitarianism and Beyond*, eds Amartya Sen and Bernard Williams (Cambridge: Cambridge University Press).

<sup>13</sup> Roemer *Theories of Redistributive Justice*, p. 144.

---

<sup>14</sup> Roemer *Theories of Redistributive Justice*, pp. 147-150 also uses a similar argument against Harsanyi's "original position theorem". Also known as the "Impartial Observer Theorem" by Weymark "A Reconsideration": "Impersonal Choice Utilitarianism" by Sen "Social Choice Theory"; "Nonteleological Utilitarianism" by Binmore, *Just Playing*, and *Playing Fair*; and the "Equiprobability Model" by Harsanyi, "Morality and the Theory of Rational Behaviour".

<sup>15</sup> See for example, Binmore, *Just Playing*, p. 526.

<sup>16</sup> Such as Roger B. Myerson, "Utilitarianism, Egalitarianism, and the Timing Effect in Social Choice Problems" *Econometrica*, 49 (1981): 883-897; or L. Epstein and U. Segal "Quadratic Social Welfare Functions" *Journal of Political Economy*, 100 (1992): 691-712.

<sup>17</sup> Amartya Sen, "Information and Invariance in Normative Choice" in *Social Choice and Public Decision-Making: Essays in Honor of Kenneth J. Arrow vol 1* eds Walter P. Heller, Ross M. Starr and David A. Starett (Cambridge: Cambridge University Press, 1986), p. 40.

<sup>18</sup> See Yew-Kwang Ng "Happiness Surveys: Some Comparability Issues and an Exploratory Survey Based on Just Perceivable Increments" *Social Indicators Research*, 38 (1996): 1-27 by someone in favour in representing utility in terms of "happiness".

<sup>19</sup> Sen, "Information and Invariance." p. 40.

---

<sup>20</sup> Keith Dowding, “Revealed Preference and External Reference” *Rationality and Society*, 14 (2002): 257-282.

<sup>21</sup> Note here that “desire” is not “self-interested desire”.

<sup>22</sup> Peter A. Diamond, “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility: A Comment” *Journal of Political Economy*, 75 (1967): 765-766.

<sup>23</sup> Harsanyi, “Cardinal Welfare, Individualistic Ethics”. Though we may note that Myerson, “Utilitarianism, Egalitarianism and the Timing Effect” demonstrates Harsanyi’s theorem without using the sure-thing principle. Here he assumes the social choice function is linear, that is everyone can expect the same utility from a social choice whether it is made in advance of its outcomes, or concurrently with them.

<sup>24</sup> R. Deschamps and L. Gevers, “Separability, Risk-Bearing and Social Welfare Judgements” in *Aggregation and the Revelation of Preferences* ed. J-J Lafont (Amsterdam: North-Holland, 1979): 145-160, pp. 149-150.

<sup>25</sup> John Broome, “Uncertainty and Fairness” *Economic Journal*, 94 (1984): 624-632, p. 627.

<sup>26</sup> Broome, “Uncertainty and Fairness” p. 627.

<sup>27</sup> There is a whole literature attacking “consequentialism” (or standard decision theory) along these lines.

---

<sup>28</sup> Amartya Sen *Collective Choice and Social Welfare* (Edinburgh: Oliver and Boyd, 1970), pp. 144-145.

<sup>29</sup> Amartya Sen, *Collective Choice and Social Welfare* (Edinburgh: Oliver and Boyd, 1970) ch 6 and 6\*, “The Impossibility of a Paretian Liberal” in *Choice, Welfare and Measurement*, (Oxford: Blackwell, 1982): 285-290.

<sup>30</sup> Keith Dowding and Bruno Verbeek “The Impossibility of Paretian Welfarism,” Paper presented at the European Public Choice Conference, Paris April 18-21, 2000.

<sup>31</sup> Amartya Sen, “Foundations of Social Choice Theory: An Epilogue” in *Foundations of Social Choice Theory*, eds Jon Elster and Aanund Hylland (Cambridge: Cambridge University Press, 1986): 213-248.

<sup>32</sup> Sen, “Utilitarianism and Welfarism,” and “Informational Analysis”. See also Amartya Sen, “Welfare, Freedom and Social Choice: A Reply” *Recherches Economiques de Louvain* 56 (1990): 451-485. C. D. Blackorby, D. Donaldson, and J. Weymark, “A Welfarist Proof of Arrow's Theorem” *Recherches Economiques de Louvain* 56 (1990).

<sup>33</sup> Robert E. Goodin, “Laundering Preferences” in *Foundations of Social Choice Theory*: 75-101; Sen “Foundations”, p. 222.

<sup>34</sup> Jeremy Bentham “Article on Utilitarianism (Long Version)” in *Deontology Together with A Table of the Springs of Action and Article on Utilitarianism*, ed. Amnon Goldworth (Oxford: Clarendon Press, 1983), pp. 304-5: “By the *ipsedixit*

---

principle, understand that principle or say that commencement or train of reasoning .. [that] either tacitly or expressly and avowedly the opinion - the declared opinion - of either the writer or speaker himself or some other individual named or unnamed.”

<sup>35</sup> I apologize for some of the silliness of this discussion but in my view, silly examples breed silly discussions. See Daniel C. Dennett, *Elbow Room: the Varieties of Free Will Worth Wanting* (Cambridge, Mass.: MIT Press, 1984), esp ch 2 for a discussion of how to treat “intuition pumps” in philosophy.

<sup>36</sup> Robert E. Goodin, *Green Political Theory* (Cambridge: Polity Press, 1992).

<sup>37</sup> Brian Barry, *Justice as Impartiality* (Oxford: Oxford University Press, 1995), p. 22.

<sup>38</sup> Binmore *Just Playing*, pp. 358-359.

<sup>39</sup> Binmore, *Just Playing*, pp. 358.

<sup>40</sup> Another way of understanding this is that if the consistency requirements of revealed preference seem to be broken when we observe a person’s behaviour then we have not correctly understood their utility function. We have not interpreted their actions correctly. In this sense revealed preference analysis is diagnostic, see Dowding, “Revealed Preference and External Reference”.

<sup>41</sup> This is rather too simple. Binmore uses the term “tastes” provocatively, but even tastes proper can be explained by reasons. I believe the reason I do not like sauvignon blanc is because of the occasion as a child when I picked and ate too many

---

gooseberries from the garden of friend and made myself sick. Similarly many tastes can be based on fashion itself based on commercial needs, see the discussion of preferences for white and brown bread in Keith Dowding, *Rational Choice and Political Power* (Aldershot: Edward Elgar, 1991), p. 32ff.

<sup>42</sup> Note this requires that the theory generates an equilibrium. We may also note that rarely, if ever, do philosophers actually do this. Generally they provide moral theories which justify the moral intuitions we or they already have.

<sup>43</sup> John Broome, “Utility” and “A Reply to Sen” *Economics and Philosophy* 7 (1992): 1-12, 285-287; Don Ross, *The Concept of Utility from Bentham to Game to Theory* (Cape Town: University of Cape Town Press, 1999).

<sup>44</sup> Allan Gibbard, “Ordinal Utilitarianism” in *Arrow and the Foundations of the Theory of Economic Policy*, ed. G. R. Feiwel (London: Macmillan, 1987); David Sobel, “Well-Being as the Object of Moral Consideration” *Economics and Philosophy* 14 (1998): 249-281.

<sup>45</sup>. In fact people tend to “over” punish, that is set up inefficient punishment schemes, see for example Elinor Ostrom, Roy Gardner and James Walker, *Rules, Games, and Common-Pool Resources* (Ann Arbor: Michigan University Press, 1994), ch. 8.

<sup>46</sup> Ronald Dworkin, *Taking Rights Seriously* (Cambridge, Mass.: Harvard University, 1977).

---

<sup>47</sup> This is probably the only occasion that the statement “that is part of her culture” is a legitimate explanation (and potential critique) of someone’s behaviour. That is, when the practice can only be justified on grounds that no longer hold, so the practice is continued only through replication.

<sup>48</sup> Richard Dawkins, *The Selfish Gene* (Oxford: Oxford University Press, 1976); *The Extended Phenotype* (Oxford: Oxford University Press, 1999).