

## SYMPOSIUM ON AMARTYA SEN'S PHILOSOPHY: 5 ADAPTIVE PREFERENCES AND WOMEN'S OPTIONS

MARTHA C. NUSSBAUM

*University of Chicago Law School*

---

The doctor was rightly upset about [the unsanitary conditions in the women's quarters]; but he was wrong in one respect. He thought that it was a source of constant pain for us. Quite the contrary . . . To those with low self-regard, neglect does not seem unjust, and so it does not cause them pain. That is why women feel ashamed to be upset about the injustice they encounter. If a woman must accept so much injustice in the life ordered for her, then it is perhaps less painful for her to be kept in total neglect; otherwise, she is bound to suffer, and suffer pointlessly, the pain of injustice, if she cannot change the rules governing her life. Whatever the condition that you kept us in, it rarely occurred to me that there was pain and deprivation in it. (Rabindranath Tagore, 'Letter from a Wife' (1914))

When we make videos, and women like us watch them, we get confidence to try and make changes. When we see women like us who have done

This paper is based on Chapter 2 of *Women and Human Development: The Capabilities Approach (WHD)*. The book develops an approach to the foundations of basic political principles using an idea of human capability and functioning; for the list of the Central Capabilities, see Appendix. I defend the use of the capabilities list (together with the idea of a threshold that determines a basic social minimum) as the ethical core of a form of political liberalism in the Rawlsian sense, that is, a core of judgements endorsed for specifically political purposes by people who otherwise have different comprehensive conceptions of the good. In the Introduction I discuss differences between my version of the capabilities approach and Sen's.

Vasanti and Jayamma are poor working women from Ahmedabad, Gujarat, and Trivandrum, Kerala, respectively; I met and talked with them during a field trip. Kokila is a friend of Vasanti's who helps her report cases of domestic violence to the police. Their stories are told more fully in *WHD*, Introduction.

something brave and new, then we get the confidence that we can learn something new too. When poor women see other poor women as health workers on the video, they say, 'I can also learn about health and help solve these problems in my neighborhood'. When other self-employed women see me, a vegetable vendor, making these films, they also have the confidence that they can do things which at first seem impossible. (Lila Datania, SEWA (Self-Employed Women's Organization), Ahmedabad, 1992).

### 1. PREFERENCE AND THE GOOD: TWO UNSATISFACTORY EXTREMES

Any defense of universal norms involves drawing distinctions among the many things people actually desire. If it is to have any content at all, it will say that some objects of desire are more central than others for political purposes, more indispensable to a human being's quality of life. Any wise such approach will go even further, holding that some existing preferences are actually bad bases for social policy. The list of Central Human Capabilities that forms the core of my political project contains many functions that many people over the ages have preferred not to grant to women, either not at all, or not on a basis of equality. To insist on their centrality is thus to go against preferences that have considerable depth and breadth in traditions of male power. Moreover, the list contains many items that women over the ages have not wanted for themselves, and some that even today many women do not pursue – so in putting the list at the center of a normative political project aimed at providing the philosophical underpinning for basic political principles, we are going against not just other people's preferences *about* women, but, more controversially, against many preferences (or so it seems) *of* women about themselves and their lives. To some extent, my approach, like Sen's, avoids these problems of paternalism by insisting that the political goal is capability, not actual functioning, and by dwelling on the central importance of choice as a good. But the notion of choice and practical reason used in the list is a normative notion, emphasizing the critical activity of reason in a way that does not reflect the actual use of reason in many lives.

Think of Vasanti and Jayamma. Vasanti stayed for years in an abusive marriage. Eventually she did leave, and by now she has very firm views about the importance of her bodily integrity: indeed, she and Kokila spend a lot of their time helping other battered women report their cases to the police and goading the police to do something about the problem. But there was a time when Vasanti did not think this way – especially before her husband's vasectomy, when she thought she might still have children. Like many women, she seems to have thought that abuse was painful and bad, but, still, a part of women's lot in life, just something women have to put up with as part of being a woman

dependent on men, and entailed by having left her own family to move into a husband's home. The idea that it was a violation of rights, of law, of justice, and that *she herself* had rights that were being violated by his conduct – she did not have these ideas at that time, and many many women all over the world still do not have them. My universalist approach seems to entail that there is something wrong with the preference (if that is what we should call it) to put up with abuse, that it just should not have the same role in social policy as the preference to protect and defend one's bodily integrity. It also entails that there is something wrong with not seeing oneself in a certain way, as a bearer of rights and a citizen whose dignity and worth are equal to that of others.

Or consider Jayamma, a great defender of her bodily integrity, but very acquiescent in a discriminatory wage structure and a discriminatory system of family income sharing. When women got paid less for heavier work in the brick kiln and denied chances of promotion, Jayamma did not complain or protest. She knew that this was how things were and would be. Like Tagore's character in my epigraph, she did not even waste mental energy getting upset, since these things could not be changed. Again, when her husband took his earnings and spent them on himself in somewhat unthrifty ways, leaving Jayamma to support the children financially through her labor, as well as doing all of the housework, this did not strike her as wrong or bad, it was just the way things were, and she did not waste time yearning for another way. Unlike Vasanti, Jayamma seemed to lack not only the concept of herself as a person with rights that could be violated, but lacked, as well, the sense that what was happening to her was a wrong.

Finally, let me introduce one further example, to show the way entrenched preferences can clash with universal norms even at the level of basic nutrition and health. In the semi-arid area outside Mahabubnagar, Andhra Pradesh, I talked with women who were severely malnourished, and whose village had no reliable clean water supply. Before the arrival of a government consciousness-raising program, these women apparently had no feeling of anger or protest about their physical situation. They knew no other way. They did not consider their conditions unhealthy or unsanitary, and they did not consider themselves to be malnourished. Now their level of discontent has gone way up: they protest to the local government, asking for clean water, for electricity, for a health visitor. They cover their food supplies from flies, they wash their bodies more often. Asked what was the biggest change that the government program had brought to their lives, they immediately said, as if in chorus, 'We are cleaner now'. The consciousness-raising program has clearly challenged entrenched preferences and satisfactions, taking a normative approach based on an idea of good human functioning.

Feminists who challenge entrenched satisfactions are frequently charged with being totalitarian and anti-democratic for just this way of proceeding. Who are they to tell real women what is good for them, or to march into an area shaped by tradition and custom with universal standards of what one should demand and what one should desire?

The capabilities approach, as I develop it, uses an account of central human capabilities to ground basic political principles that can be embodied in constitutional guarantees. Economists and others who defend preference-based views rarely make a clear distinction between their use in social choice generally and their use in selecting such basic principles. Instead, they tend to make general pronouncements about social choice. Insofar as they do so, I shall simply assume that they make no exception for the special situation where one is selecting basic political principles, and my critique will be focused on this special situation, rather than other situations in which appeals to preference might well play a valuable role. I think this is fair, because I think that my opponents intend their position as a perfectly general account of social choice.

In the debate about how preferences should figure in social choice, we can identify two extreme positions, between which I shall situate my own. The first position can be called *subjective welfarism*. This position holds that all existing preferences are on a par for political purposes, and that social choice should be based on some sort of aggregation of all of them. The second position can be called *platonism*.<sup>1</sup> According to this view, the fact that people desire or prefer something is basically not relevant, given our knowledge of how unreliable desires and preferences are as a guide to what is really just and good. Actual desire and choice play no role at all in justifying something as good. What we need to do is to provide an argument for the objective value of the relevant state of affairs that is independent of the fact that people desire or prefer it; once we have such an argument, we are justified in making even radical departures from people's actual wants.

Both positions are motivated by genuinely important concerns. Welfarism springs from respect for people and their actual choices, from a reluctance to impose something alien upon them, or even to treat the desires of different people unequally. In effect, it starts from respect for persons, interpreting that as equivalent to respect for preferences. Platonism springs from an urgent concern for justice and human value, and from the recognition that in the real world these values are frequently subordinated to power, greed, and selfish indulgence. But both contain obvious problems. Embraced as a normative position, subjective welfarism makes it impossible to conduct a radical critique of

<sup>1</sup> This may or may not be a position held by Plato.

unjust institutions; it forces us to say, for example, that because Jayamma has accepted an unjust wage structure as the way things must be, that is the way they should remain. This limitation is especially grave when we are in the process of selecting basic political principles that can be embodied in constitutional guarantees. Platonism, on the other hand, seems too disdainful of the wisdom embodied in people's actual experience: it seems to care too little about what Jayamma and Vasanti think about changes in their lives. Any viable modern position, it seems, must try to preserve the important values contained in each of these two extremes, while avoiding their defects.

In order to search for an appropriate 'mean' between these extremes, it will be useful to look, first, at why utilitarian economists have rejected pure subjective welfarism, and how they have tried to refurbish the welfarist view. Next, I shall look at some reasons we might have for concluding that the problem of preference-deformation requires us to depart altogether from the utilitarian framework. Finally, I shall show how my own view of the central capabilities addresses these issues, arguing that at the level of the central capabilities there is considerable convergence between a 'substantive good' approach to social goods and an 'informed desire' approach.<sup>2</sup>

## 2. WELFARISM: THE INTERNAL CRITIQUE<sup>3</sup>

Although subjective welfarism is common in brief gestures toward a normative approach in economics, very few utilitarian economists have been willing to be thoroughgoing subjective welfarists, once they consider normative issues head on and extensively. Milton Friedman (1984, p. 210) probably was the real article: he clearly did hold that concerning differences of value 'men can ultimately only fight', and that, in consequence, there was nowhere for normative theory to go beyond subjective welfarism. Usually, however, economists have recognized that it is implausible to treat all existing preferences as on a par for normative purposes, and have recommended at least some winnowing or correcting.

The most obvious such correction involves false belief and lack of information. Even Hume, who in general thought that passions and

<sup>2</sup> For these terms, see Thomas Scanlon (1993, pp. 201–7). Scanlon actually uses the term 'substantive list', but also notes that the term is perilous, suggesting that we are talking about a laundry list of unrelated items, rather than a coherent view.

<sup>3</sup> I omit at this point a preliminary section in which I dissect the concept of 'preference' typically used in economic discussions, pointing out that it is both contested and quite vague; the relevant arguments can be found in *WHD* Chapter 2 and at greater length in Nussbaum (1997, pp. 1197–214).

desires could not be coherently deemed 'unreasonable', made exceptions for cases in which one mistakenly believes an object to exist that does not exist, or holds false beliefs about appropriate means to further ends.<sup>4</sup> But most utilitarian followers of Hume go somewhat further than he did in the recognition of cognitive error. To take a representative example, Christopher Bliss (1993, pp. 418–19), while defending subjective welfarism in a very strong form in connection with the measurement of quality of life in developing countries, recognizes the need to correct for inadequate or false information: 'The inhabitants of a poor country for example, may not realize how unhealthy they are, and the consequences of that ill health, whereas an expert will know'. Another case in which Bliss would admit expert corrections of existing views is the case where we need a global overview and individuals are unable to provide this. Again, holds Bliss (1993, p. 419), this correction is acceptable because it involves 'imperfect vision' on the part of the individual. We can correct the individual's vision – give her eyeglasses, so to speak – without losing hold of 'the fundamental point that man, if not the measure of all things, is at least the measure of the standard of living'. Bliss does not discuss the grounding of basic political principles that underlie constitutional guarantees, so it is possible that he would not endorse a preference-based view in that domain; on the other hand, his zealous defense of preferences in all areas of human life quality gives no suggestion that he sees any area in which reliance on preferences will prove problematic.

A much more ambitious set of corrections to existing preferences was proposed by John Harsanyi (1982, pp. 39–62), still apparently within the general framework of subjective welfarist theory. Harsanyi (1982, p. 55) begins by announcing 'the important philosophical principle of *preference autonomy*. By this I mean the principle that, in deciding what is good and what is bad for a given individual, the ultimate criterion can only be his own wants and his own preferences'. (Harsanyi does not clearly explain his reasons for holding this principle, but it would appear that, in addition to the usual concerns about democracy, he is also motivated by the thought that we simply cannot make sense of the idea that what A wants is bad for A, except as a claim that, at some deeper level, A really prefers something else. (p. 55)) On the other hand, Harsanyi recognizes that people's preferences are frequently 'irrational'. He believes that 'any sensible ethical theory' must recognize this fact. 'It would be absurd to assert that we have the same moral obligation to help other people in satisfying their utterly unreasonable wants as we have to help them in satisfying their very reasonable desires' (p. 55). But

<sup>4</sup> Hume, *A Treatise of Human Nature*, Bk. II, Pt. III, Sec. 3, concluding with the famous judgement that 'It is not contrary to reason to prefer the destruction of the whole world to the scratching of my finger'.

what content can we give to this distinction, compatibly with maintaining the welfarist principle? A normative hedonist, he observes, could easily make the relevant distinction: a rational want will be a want for something that really does produce pleasure, and an irrational want will be for something that does not really produce pleasure. But if we do not accept that sort of definite normative theory (and Harsanyi has already rejected it, apparently in favor of a theory that derives normativity from preferences themselves), then what are we to say?

We must say, Harsanyi concludes, that people's manifest preferences are frequently at odds with their 'true preferences'. A person's rational wants are those that are consistent with his true preferences, and the irrational are those that are not. The distinction between the manifest and the true is defined as follows:

[A person's] manifest preferences are his actual preferences as manifested by his observed behaviour, including preferences possibly based on erroneous factual beliefs, or on careless rational choice. In contrast, a person's true preferences are the preferences he *would* have if he had all the relevant factual information, always reasoned with the greatest possible care, and were in a state of mind most conducive to rational choice. (1982, p. 55)

Social utility, he then concludes, should be defined in terms of the true, rather than the manifest, preferences of individuals, and the maximization of social utility is the appropriate social goal. Harsanyi puts this idea forward as a perfectly general theory of social choice, especially in the area of fundamental principles: he characterizes his preference-based ethical theory as an alternative to Rawls's theory of justice.<sup>5</sup>

To come up with something that seems even *prima facie* satisfactory, notice, Harsanyi has had to add not only the usual corrections to belief and information, but also the strongly normative procedural idea of careful reasoning and the 'state of mind most conducive to rational choice'. He does not spell out this last ingredient further, but when we think of our cases we can easily spot some people who are *not* in a state of mind that seems conducive to rational choice: Vasanti, intimidated by her husband's physical abuse and terrified about her survival prospects should she leave him; Jayamma, habituated to thinking that unequal control over household income is just women's lot; Tagore's character Mrinal in my epigraph, used to thinking of herself as of low worth. Those conditions certainly do not strike us as conducive to rational choice, and that is just the point Tagore's character is making, as she explains to her husband how she has finally decided to leave him. For her, adequate choice-making required, first of all, throwing off the

<sup>5</sup> Harsanyi (1982, pp. 40–41).

slumberous state induced by years of contempt and neglect. But if we were to put absence of traditional hierarchy, absence of fear, and a sense of one's worth and dignity into the rational choice process, we would be moving very far indeed from a standard welfarist approach. I shall suggest later that these additions really do help us construct an informed-desire approach that is of some heuristic value; but it is certainly quite unclear whether Harsanyi means to take us so far from his welfarist starting point.

Again, consider the women of SEWA in my second epigraph, who see videos of women doing daring new things and thereby get confidence that they can do these things too. Now clearly it is Lila Datania's point that the experience of watching the videos helps these women make adequate choices for their future – not only by giving them new information but by enhancing their sense of their possibilities and their worth. But we would not think this progress, or a correction of malformed preferences in the direction of 'true' preferences, if the women were taught by the videos to hide away in the house all day, or to believe that they were made for physical abuse. And yet we know well that videos (violent pornographic videos, for example) can teach people such attitudes about themselves and others. It is because we have an implicit theory of value that holds self-respect and economic agency to be important goods that we think the preferences constructed by the videos are good; it is not clear that there would be any purely formal way to make the distinction.

Harsanyi makes one further correction to welfarism that takes him more clearly away from welfarism. This is, that some people's 'true' preferences will have to be excluded altogether from the social-utility function: 'In particular, we must exclude all clearly antisocial preferences, such as sadism, envy, resentment, and malice'.<sup>6</sup> Harsanyi's justification for this move is fascinating for the way in which it reveals moral intuitions of a non-welfarist kind that lie beneath his welfarism. The 'fundamental basis' of our moral commitments to others in utilitarianism, he says, is 'a general goodwill and human sympathy'. Thus '[u]tilitarian ethics makes all of us members of the same moral community'. But if this idea of a moral community is what lies behind utilitarianism and gives its normative judgements their appeal, then we must interpret utilitarianism to demand the exclusion of parts of real people's personalities that are hostile to the idea of a moral community: 'A person displaying ill will toward others does remain a member of this community, but not with his whole personality. That part of his personality that harbours these hostile antisocial feelings must be

<sup>6</sup> 1982, p. 56, taking issue with J. J. C. Smart's version of normative utilitarian theory.

excluded from membership, and has no claim for a hearing when it comes to defining our concept of social utility' (1982, p. 56).

What is fascinating here is to see that an idea of quite a Kantian sort lies, for Harsanyi, at the bottom of the utilitarian social choice function<sup>7</sup> – and it is only as regulated by that ideal vision of a community of ends that the utility function can prove acceptable as a basis of social policy. Harsanyi does not say enough here for us to ascribe to him a very definite ethical conception. But his real interest in preferences turns out to be, at bottom, something like Kantian interest in respecting persons, their equality and their autonomy, and that he is not really averse to any departure from utilitarianism that preserves these essential Kantian features. In short, his view is only in appearance a welfarist view at all.

One more outpost on the road away from welfarism should now be considered, since it is probably the smartest and most consistent attempt to refurbish the preference-based view of social choice. This is Richard Brandt's view of 'cognitive psychotherapy' in *A Theory of the Good and the Right*.<sup>8</sup> Brandt, like Harsanyi (or the apparent Harsanyi) is, ultimately, a welfarist. He holds that the ultimate criterion of both personal and social rationality must be found within each person, and not by importing any values external to that person's own values. But since Brandt recognizes that errors are frequently deeply implanted in people and cannot always be driven out by a simple disclosure of the correct facts, he concludes that we can get to the person's true preferences only by a prolonged process of 'cognitive psychotherapy', carefully defined as 'value-free reflection' that 'relies simply upon reflection on available information, without influence by prestige of someone, use of evaluative language, extrinsic reward or punishment, or use of artificially induced feeling-states like relaxation' (p. 113). This process is then used to define rationality in desire: 'I shall call a person's desire, aversion, or pleasure "rational" if it would survive or be produced by careful "cognitive psychotherapy" for that person. I shall call a desire "irrational" if it cannot survive compatibly with clear and repeated judgements about established facts'. We notice already that the absence of authority, intimidation, and hierarchy in the method is itself not so clearly value-neutral; it expresses values – independence, liberty, self-driven choice – that Brandt actually thinks very important. These are indeed important values to build into a procedure of cognitive scrutiny of desire. What is questionable is that we should think of the resulting method as entirely value-free.

<sup>7</sup> He begins his article by announcing a debt to Kant and Adam Smith, as well as to the Utilitarians (1982, pp. 39–40).

<sup>8</sup> Richard B. Brandt (1979, pp. 234–45). By calling his view an alternative to John Rawls's theory of justice, Brandt makes it clear that he means to use it to choose basic political principles.

Brandt now identifies four categories of mistake that cognitive psychotherapy would, in his view, remove (pp. 115–26). First, there is the large category of desires that depend upon false beliefs (and recall that this must not include beliefs about matters of value, which remain unaffected by the cognitive process). Second, there are generalizations from atypical examples; these, too, can ultimately be dislodged by a more extensive confrontation with a wider range of examples. Third,<sup>9</sup> there is the category called ‘artificial Desire-Arousal in Culture-Transmission’. What Brandt means here is that cultures transmit values by example and precept, frequently in ways that could not have been produced by the real experience of what is talked about, without cultural interference. Brandt’s two examples are the occupation of garbage collecting, and ‘marriage to a person of another race, religion, or nationality’. Actual experience with these activities, says Brandt, could prove very satisfying, and certainly there’s no reason to suppose that they would produce an intrinsic aversion in someone who had had no prior cultural conditioning. Of course, he now says, the social attitudes of others are themselves real facts with which people have to deal. ‘But intense concern with the attitudes of other people is itself founded upon error – the false belief that the attitudes of others are crucially important for an adult, especially if the attitudes in question are those of one’s own parents only’.

In this fascinating paragraph we see Brandt trying to the utmost to squeeze conclusions that please him out of the allegedly value-free method that he recommends. Brandt’s love of liberty, his democratic respect for those who do manual labor, and his dislike of superstition all make him see the attitudes of those who shrink from intermarriage and garbage collecting as profoundly irrational. Probably he is right to say that an untutored child would experience no natural disdain for these things, but it is not clear that this is the line he should take, since an untutored child might also lack the basis for many social attitudes Brandt does not want to get rid of, such as an aversion to cruelty, a concern for the well-being of the poor, a love of free speech, a passion for justice. So the value-free science removes too much, if it really removes all attitudes that require evaluative learning in a culture for their transmission. As for the allegedly value-free false belief that the attitudes of others, including one’s parents, are unimportant, once again, this is so only given a certain scheme of ends and values, and not given others. In short, Brandt’s liberal and democratic instincts clash, as did Harsanyi’s, with what his argument can actually deliver – despite the fact that he is more hard-headed than Harsanyi in his attempts to do without values.

<sup>9</sup> I have inverted the order of categories two and three, since I find Brandt’s original third straightforward and his original second very problematic.

Brandt's fourth category of mistake is that of 'Exaggerated Valences Produced by Early Deprivation' (1979, pp. 122–6). His example is a letter to Dear Abby, in which a woman complains about her husband, who grew up fatherless during the depression, and has now become quite wealthy. Nonetheless, he is obsessed with saving for his old age, to a degree that makes his family unhappy: he buys second-hand clothes, eats stale bread, and so on. (Brandt notes that similar behavior can be found in laboratory rats.<sup>10</sup>) Here, he says, we have a syndrome in which early deprivation and its associated anxieties lead to an exaggerated later development of desire. He claims that such abnormally strong desires would diminish once their root cause was brought to the surface – both in the case of money and in the case of other goods, such as love or affection, that might also be lacking in one's early life.

But of course the unresolved question is, when is a desire for money, or for love, 'exaggerated?' Are we supposed to be able to tell that without a theory of value? And what counts as 'deprivation' in early life? Again, we need a normative account of proper love, and proper material support, to get started here. By choosing a very bizarre case, Brandt conceals the depth of this problem. But here again, clearly, he is getting to some characteristic Brandtian value-conclusions (for example, people should not be very dependent on the love and approval of others) through an allegedly factual exercise.

In short, the welfarist attempt to refurbish the preference- or desire-based view runs into difficulty; it appears unable to deliver all that the welfarist philosopher-economists themselves would like to get. We can get a certain distance by adding information and correcting logical error. But to get all the way to Harsanyi's ideal of a moral community of equals, or to Brandt's ideal of an independent hard-headed unsuperstitious democratic citizenry, these thinkers have had to inject value judgements into the procedures of revision – contrary to their announced intentions.

### 3. ADAPTIVE PREFERENCES AND THE REJECTION OF WELFARISM

It is not surprising, when zealous defenders of the welfarist project still find themselves diverging from it, that others less wedded to the project should conclude that, for normative purposes – and especially for the purposes of choosing basic political principles – the informed-desire approach is inadequate. In recent years there has been an explosion of work attacking the preference-based approach to normative issues of public choice. But it is worth reviewing the different arguments that

<sup>10</sup> p. 123: rats who have been deprived of food respond by massive hoarding when food becomes abundant.

have been made, in order to see exactly how far they do break with a preference-based view, and in what ways.

1. *The Argument from Appropriate Procedure.* The informed-desire approach struck even Harsanyi and Brandt as in need of procedural supplementation: in different ways, they each built into the procedure the idea of a community of equals, unintimidated by power or authority, and unaffected by their own awareness of their place in a social hierarchy, whether for envy or for fear. And this of course has been a tremendous area of normative work, among thinkers who continue to believe that some form of proceduralism will suffice as a basis for social choice. The views of John Rawls and Jürgen Habermas about fair procedures of public choice are both too familiar and too complex for me to get involved in discussing them here, but both clearly do with rigor and detail what Harsanyi did not fully do, but only mentioned: that is, to model a Kantian ideal of moral community, by introducing constraints on information and procedure. To this distinguished list we should add Jean Hampton (1993, pp. 227–56), whose feminist proceduralism, in the important article ‘Feminist Contractarianism’, builds into the procedure of choice the value of each person as an end, and a distinction between legitimate and illegitimate interests.

2. *The Argument from Adaptation.* Closely linked to these normative criticisms of utilitarianism is a set of arguments that focus on the phenomenon of *adaptation*, in which individuals adjust their desires to the way of life they know. Jon Elster’s account of adaptation is rather narrowly focused.<sup>11</sup> For him, a desire counts as adaptive only if it really has a fox-and-grapes structure: having desired the grapes, the fox, seeing that he cannot get the grapes, judges that they are sour. Such preferences are to be distinguished from changes in desire based on learning and experience: for the latter are likely to be irreversible, whereas adaptive preferences (for city life when in the city, for country life when in the country) are far from irreversible.

Elster’s somewhat romantic preference for striving and yearning makes him suspicious of any desire that is formed through adjustment to reality. But it is not at all clear he should in such a sweeping way condemn adaptive preferences. We get used to having the bodies we do have, and even if, as children, we wanted to fly like birds, we simply drop that after a while, and are the better for it. Again, someone as a child may want to be the best opera singer in the world (as I did), or the best basketball player – and most people adjust their aspirations to what they can actually achieve. It seems that these changes do involve the fox-

<sup>11</sup> In ‘Sour Grapes’, in *Utilitarianism and Beyond*, pp. 219–38, and in the book of the same title. Subsequent page references are to the article. *WHD* contains a more detailed discussion of Elster’s position.

and-grapes structure that Elster rules out as incompatible with autonomy: they are adjustments in response to a perception of one's circumstances, rather than the result of deliberate character formation, and they lack the condition of 'freedom to do otherwise' that Elster introduces as a necessary condition for autonomous wants, to distinguish them from adaptive wants (pp. 228–9). I am not free to be a leading opera singer, nor is the short adult free to be a leading basketball player. We have failed to reach the grapes and we have shifted our preferences in keeping with that failure, judging that such lives are not for us. But clearly this is often a good thing, and we probably should not encourage people to persist in unrealistic aspirations.<sup>12</sup>

The cases that Elster actually has in mind are interestingly different: the way in which feudalism made people not aspire to political equality and material well-being, the way the Industrial Revolution unleashed a storm of class-based discontent that was productive both politically and economically. But to distinguish this case from my cases of the bird and the basketball player, he needs something he does not give us, a substantive theory of justice and central goods. It was fruitful for these people to hold onto their (*pro tempore*) unrealizable desires because they were desires for central goods, things people as people have a right to have. People's liberty can indeed be measured, not by the sheer number of unrealizable wants they have, but by the extent to which they want what human beings have a right to have. Thus Vasanti, who hated her domestic abuse, seems a little more free than Jayamma, who acquiesced in discrimination and oppression. Both, however, were unfree in one further crucial way: they lacked the understanding of themselves as citizens who have rights that are being violated. That type of adjustment to bad circumstances is indeed deplorable, and we view it as progress that they come to realize they have a right to better treatment, even if that better treatment is not yet forthcoming. But to say this, we need an account of what types of treatment people have a right to expect in central areas of their lives. Once again, proceduralism – even of a more complicated sort – seems insufficient without something in the way of a substantive theory.

Such a theory is to some extent provided by the other prominent economist who discusses adaptive preferences, Amartya Sen.<sup>13</sup> Sen focuses on the situation of women and other deprived people; his central case is that of women who do not desire some basic human good because they have been long habituated to its absence or told that it is not for such as them. For example, in 1944, the year after the Great

<sup>12</sup> Contrast Elster, p. 228, holding that a person's degree of autonomy can be measured by the number of things he wants to do but is not at liberty to do.

<sup>13</sup> See, for example, Sen 1995, pp. 259–73; 1984, pp. 307–24.

Bengal Famine, the All-India Institute of Hygiene and Public Health did a survey in an area near Calcutta, including in the survey many widows and widowers. Among the widowers, 45.6 percent ranked their health as either 'ill' or 'indifferent'. Only 2.5% of widows made that judgement, and none at all ranked their health as 'indifferent' (as Sen notes, a more subjective category than 'ill'). This was in striking contrast to their real situation, since widows tend to be a particularly deprived group in basic health and nutrition. Sen concludes: 'Quiet acceptance of deprivation and bad fate affects the scale of dissatisfaction generated, and the utilitarian calculus gives sanctity to that distortion'. One can also make a remark in the other direction: privileged people get used to being pampered and cared for, and may feel an unusually high level of discontent when the one that did the pampering is no longer around. Sen concludes that this makes utility quite inadequate as a basis of social choice.

Sen's group of cases is, notice, both broader and narrower than Elster's. It is broader, because Sen includes life-long habituation, and does not focus simply on giving up a desire one once had. And this is important where women are concerned, since most of the interesting cases do involve life-long socialization and absence of information. The group is narrower because the cases on which he dwells all involve a central human capability. Although Sen has never been willing to endorse a substantive theory of the central capabilities, in practice he does so, and therefore does not trouble himself about the adaptive preference of someone who gives up his dream of basketball stardom when he finds that he is never going to be taller than five feet four inches (and let us suppose he is not Muggsy Bogues<sup>14</sup>). There is no romantic preference for striving as a good in itself in Sen's view; the appropriateness of desire is tethered, implicitly at any rate, to a sense of basic goods of life. Sen's analysis of adaptation corresponds well to what we find in the cases of Jayamma, Vasanti, and the women of Andhra Pradesh. All, because of social habituation, undervalue central human capabilities that they later come to value.

Finally, the phenomenon of adaptive preferences was discussed in a particularly illuminating way by John Stuart Mill in *The Subjection of Women*. Mill argues that men maintain their power over women by shaping women's preferences and desires, or, as he puts it, 'enslav[ing] their minds'. They teach women that timidity and 'resignation of all individual will' are 'an essential part of sexual attractiveness'. Women internalize this teaching, just as men become accustomed to being 'the masters of women' (1869, pp. 15–16).

<sup>14</sup> Muggsy Bogues is an extraordinary 5'4" athlete who has played for various teams in the NBA.

How does Mill, a Utilitarian, criticize these adaptive preferences? Clearly, with a normative theory of liberty and equality. He makes some instrumental arguments about the social good that will be done by a more thorough use of women's talents, but the central advantage to which he points is 'the advantage of having the most universal and pervading of all human relations regulated by justice instead of injustice' (p. 86).

This is hardly the occasion to conduct a probing examination of Mill's Utilitarianism and its connection with his theory of liberty. But we can at least say this, and it is revealing for our purposes. Mill does think it highly relevant that the values he defends are in some sense rooted in human desire – that people who have tried both liberty and its absence will prefer liberty, that justice is a prominent object of human striving. He supports his proposals in *On Liberty* with reference to a quite Aristotelian account of the human powers and their flourishing, referring to 'a Greek ideal of self-development', and arguing that in the absence of liberty, 'human capacities are withered and starved' (1859, pp. 71, 75). But he also links this Aristotelian notion of self-development with a notion of experienced desire, saying that liberty is good in part because it satisfies certain 'permanent interests' of human beings. Like Aristotle, he regards it as more than a contingent matter that the constituent parts of flourishing are in fact powerfully and deeply desired. This does not rescue the welfarist project in its original form – but it does constrain the move to Platonism, in an appropriate way. For it would hardly be plausible to say that Vasanti's bodily integrity, or Jayamma's equality as a laborer, are things to be pursued altogether independently of their relationship to human desire and choice. The welfarist project fails, in its simplest form; but it gets something important right.

3. *The Argument from Intrinsic Worth*. Even if the welfarist can show that people desire liberty and justice – even if some modification of the welfarist procedure could be devised that reliably generated those goods (and without smuggling them somehow into the structure of the procedure itself), it would not be clear that this is the right way of justifying our social interest in these goods. In general, the failure of a person to have various basic human capabilities is important in itself, not just because the person minds it or complains about it. As Sen puts this point, 'If a person is unable to get the nourishment he or she needs, or is unable to lead a normal life due to some handicap, that failure . . . is itself important, and not made important only because he or she incurs dissatisfaction or disutility from that failure' (1984, p. 363). Another way to put this is to say that even if we could engineer things so that people were reliably adapted to a very low living standard – and, as Mill says, the 'masters of women' have in many areas done exactly that – this

would not be the end of the issue of what is good or right. These failures themselves have importance, and just the bare fact that human beings are incurring them should be enough for us.

This argument is the flip side of the adaptation argument. It tells us positively what that argument told us negatively, that we need a normative theory – preferably a theory of human capability that includes accounts of equality and liberty – to provide the normative basis that desire fails reliably to provide us.

Notice, however, that there are a number of ways of making this argument, many stopping well short of outright Platonism. The Platonist will indeed say that these eternal intrinsic values have the value they do altogether independently of human history, human choice, and human desire. But one might adopt a different account of justification, one that would make at least a qualified reference to choice and desire. Rawls's Socratic account of justification proceeding toward 'reflective equilibrium' is one such account; Hampton's feminist contractarianism, with its reference to legitimate interests, is another.

Thus we can accept the argument from intrinsic value without accepting the extreme rejection of desire urged, for example, by Thomas Scanlon in his 'Value, Desire, and the Quality of Life'.<sup>15</sup> Scanlon argues that there are only two reasons why desire is of any interest at all in the process of justifying an account of quality of life. One reason is hedonic: the satisfaction of a desire may bring pleasure, and pleasure is an intrinsic good. The other reason is heuristic: desire steers us in the direction of some intrinsic goods. But in either case, the reference of desire is dispensable: the hedonic reason points back to the intrinsic value of pleasure, which is (on this view) not valuable simply because it is desired; the heuristic reason points to items whose value must be independently arrived at. Given all this, and given that desire is frequently not such a good guide, there is no particular reason to be interested in it in constructing an account of quality of life.

I am not fully persuaded by this argument, for two reasons. First, Scanlon never asks how, in the long run, we are actually going to justify a 'substantive list' of values of the sort he prefers, without making at least some sort of reference to desire. He probably is not a thoroughgoing Platonist, but in this article he never tells us what he really is, and what the 'other grounds' are on which he would rest a judgement of desirability. Second, Scanlon fails to consider a very strong reason we have for giving desire at least some role in our process of justification: the reason of respect I have already endorsed. The fact that human beings desire something does count; it counts because we think that politics, rightly understood, comes from people and what matters to

<sup>15</sup> In *WHD* I discuss more recent writings of Scanlon's on this issue.

them, not from heavenly norms. I do not think Scanlon disagrees with this; but then I remain puzzled by the dismissive attitude he takes to informed-desire approaches.

One difficulty in assessing the relationship between Scanlon's position and my own lies in the fact that he does not explicitly commit himself to any particular analysis of the notion of desire. If one thought of desire as brutish and unintelligent, as just a mindless 'push' that moved people toward objects without involving any selectivity or intentionality, such a view would not entail Scanlon's dismissive position – one might still think we have reason to attend to the pushes that people have in their natures and to give these some weight – but it would at least explain why a basically Kantian moral position would be inclined to bypass desire in favor of something that resides in the moral domain. On the other hand, if one thinks of desire, as I do, in a more Aristotelian way, as a reaching out for 'the apparent good', and thus as involving, even at the level of appetite, a high degree of selective intentionality and responsiveness, one will have in that very picture of desire some strong reasons not to bypass it, for it seems to be a part of our humanity worthy of respect and voice.

#### 4. DESIRE AND JUSTIFICATION

In my own conception, an account of the central capabilities provides a necessary basis for political principles, giving not a complete account of the good or of human flourishing, but a political account, specifying certain capacities, liberties, and opportunities that have value in any plan of life citizens may otherwise choose. In the first part of my project I argue for such an approach using an argument based upon an intuitively powerful idea of truly human functioning, functioning that is worthy of the dignity of the human being. I claim that citizens with a wide range of comprehensive conceptions of the good can endorse this list – as a list of capabilities, not of actual functions – as a basis for getting on with life, including prominently political life. These basic goods supply politics with constraints: citizens should be provided with these, whatever else politics also pursues.

It seems to me that this sort of capabilities approach deals well with the problems that plagued the preference-based approach. It does not waste time trying to smuggle a substantive account of central capabilities into a procedure for winnowing desire: it goes directly and forthrightly to the good (and the right<sup>16</sup>), taking an unambiguously clear stand on the need for these items, as an enabling core of whatever else human beings choose. It addresses the problem of adaptive preference, again, by

<sup>16</sup> Basic liberties and opportunities, and the dignity and equality of persons, are all included in the account of the basic capabilities.

substantive rather than formal devices, as seemed necessary. A habituated preference not to have an item on the list (political liberties, literacy, equal political rights, or whatever) will not count in the social choice function, and the equally habituated preference to have these things will count. Finally, the list does justice to the intrinsic value of the items it contains, by not subordinating them to something else, such as preference-satisfaction.

It should be apparent that the approach is not paternalistic in any typical sense, since it gives such a large place to liberty, and envisages the social goal in terms of capability, not functioning.<sup>17</sup>

But what role is played by desire in the process of justifying the list of the central capabilities? Here people's intuitions about how to proceed vary greatly: some think we only put things on a sound footing when we devise a procedure that generates the good as an output, and others (I myself), tend to think that our intuitions about the central capabilities are at least as trustworthy as our intuitions about what constitutes a good procedure. I have said so far that the capabilities view embodies an intuitively powerful idea of truly human functioning that has deep roots in many different traditions. I have used this intuitively grounded argument to justify the list and its political role. But now I must make plain what role or roles I do assign to desire in the procedure of arriving at and justifying the list.

In general, the account of political justification that I would favor lies close to the Rawlsian account of argument proceeding toward reflective equilibrium: we lay out the arguments for a given theoretical position, holding it up against the 'fixed points' in our moral intuitions, and seeing how those intuitions both test and are tested by the conceptions we examine, hoping, over time, to achieve consistency and fit in our judgements taken as a whole. My argument in the first phase of my project was envisaged as a first step in the process of reaching toward such a reflective equilibrium. Before that process would be complete (if it ever would be), we would also have to lay out other competing conceptions, compare them in detail with this one, and see on what grounds ours emerged as more choiceworthy. The present paper has taken one step in that further project, by comparing the substantive-good conception with various preference-based conceptions. Now the question should be: what part in this procedure of reaching toward reflective equilibrium is played by desire? I believe that desire plays two roles here: both an epistemic role and an ancillary role in justification.

First, it seems to me very important that people from a wide variety of cultures, coming together in conditions conducive to reflective criticism of tradition, and free from intimidation and hierarchy, should

<sup>17</sup> See *WHD*, Chapter 1 for more extensive discussion of pluralism and choice.

agree that this list is a good one, one that they would choose. Finding such areas of informed agreement is epistemically valuable, in two ways: first, it points us to areas of human expression that we might have neglected or underestimated. Second, it tells us that our intuitions about what would make a political consensus possible are on the right track. The methodology that has been used to modify the list shows this: for I have drawn both on the results of cross-cultural academic discussion and on discussions in women's groups that were themselves designed to exemplify certain values of equal dignity, non-hierarchy, and non-intimidation. In other words, I have proceeded as if it were important that there should be a substantial convergence between the substantive account and a proceduralist account, where the procedure itself is structured in accordance with certain substantive values. Informed desire plays a large role in finding a good substantive list, for epistemic reasons. What we are trying to find is something that people can live by together, generating political stability among other values. That a conception has a certain relation to informed desire is at least one part of what makes it likely to do the job we want it to do.

But the fact that informed desire plays this epistemic role in discovering that which is likely to promote political stability shows that it also plays a limited and ancillary role in justifying the political conception: for, however attractive it looks, if we cannot show that it is likely to remain stable it will be difficult to justify it. It seems right to include stability among the considerations that are important in justifying a political conception. And I argue that we cannot show that our conception is likely to remain stable without some reference to informed desire.

Our interest in adaptive preferences gives us another way of approaching this issue of stability. My examples give us many reasons to suppose that Mill is correct: the preference for the central human capabilities is not merely habitual or adaptive, but has much more the unidirectional structure of preferences formed by learning (as Elster has introduced this distinction). This gives us, again, confidence that we are on the right track in designing a political conception, where stability is concerned. We learn something about the likely stability of a consensus based on central capabilities when we note, as we do, that women who have become literate find literacy valuable and even delightful, that they report satisfaction with their new condition, and that the transition in their lives begun by literacy is not one that they would wish to reverse. The same is evidently true for health and sanitation, for learning to stand up against domestic violence, and for acquiring political liberties and capabilities: people who once learn and experience these capabilities do not want to go back, and one really cannot make them go back. The delight and satisfaction that makes people unwilling to go backwards is

a very important sign that the conception we are developing is likely to be a stable one, and that regimes that thwart central capabilities are likely to prove unstable. Thwarting permanent human interests is not a wise political strategy. And this epistemic role for desire is, at the same time once again, an ancillary justificatory role, in the sense that it is important, in order to justify our conception, to show that it can be expected to have a reasonable degree of stability.

Women do, of course, choose to return to a traditional life in the home from a life of employment outside the home. They also choose to return to veiling from non-veiling. But notice that this is a change in their mode of *functioning*, not in their level of political *capability* as citizens. To argue that the preference for the central capabilities is not unidirectional, we would need to argue that people wish to give up *choices* and *opportunities* that they have in this area, as citizens choosing basic political principles. And this is far more difficult to show. What we would need to show is that women who have experienced the full range of the central capabilities choose, with full information and without intimidation (etc.), to deny these capabilities, politically, to all women. But usually women who prefer traditional lives, after having led other lives, do not campaign for a political denial of choice to all citizens.<sup>18</sup> Even in the apparent case where experienced women do campaign on behalf of general restrictions, as seemed to be the case in Iran, it is extremely unlikely that the women foresaw and wished the highly repressive regime that ensued. That regime has not been stable, because it thwarts central capabilities.

But stability is not the only issue we should consider. I have argued that desire is an intelligent part of the human being that deserves respect in itself in any procedure of justification we would design. Thus it seems to me that it is not only on account of stability that we refer to desire, but also because we respect that aspect of human personality. It is an important part of showing such respect that we do consult people from many cultures about what (under suitably informed conditions) they would wish; and their answers to those questions, again, play an ancillary role in justifying the list. It is important, again, that we take note of the delight women experience when they achieve a greater measure of capability in the central areas of human functioning: for this delight itself helps us to justify our list as one that is respectful of their personalities. Finally, of course, desire plays a role in the fact that the goal is expressed in terms of capabilities and not actual functionings. We respect the importance of desire and preference by building into the most basic level of the account the option to pursue the goal or not to pursue it.

<sup>18</sup> See further in *WHD*, Chapter 3.

In other words, as we should expect, there is and should be a good measure of convergence between an intelligently normative proceduralism and a substantive good theory of a non-Platonist kind, sensitive to people's actual beliefs and values. The capabilities approach, rooted in a respect for desire as well as in an account of human dignity and other substantive human goods, balances these concerns well, refusing to take existing preferences as a benchmark of social policy, but refusing, as well, to dismiss utterly the psychology of imperfect human beings.

#### APPENDIX: THE CENTRAL HUMAN CAPABILITIES

**1. Life.** Being able to live to the end of a human life of normal length; not dying prematurely, or before one's life is so reduced as to be not worth living.

**2. Bodily Health.** Being able to have good health, including reproductive health; to be adequately nourished; to have adequate shelter.

**3. Bodily Integrity.** Being able to move freely from place to place; to be secure against violent assault, including sexual assault and domestic violence; having opportunities for sexual satisfaction and for choice in matters of reproduction.

**4. Senses, Imagination, and Thought.** Being able to use the senses, to imagine, think, and reason – and to do these things in a 'truly human' way, a way informed and cultivated by an adequate education, including, but by no means limited to, literacy and basic mathematical and scientific training. Being able to use imagination and thought in connection with experiencing and producing works and events of one's own choice, religious, literary, musical, and so forth. Being able to use one's mind in ways protected by guarantees of freedom of expression with respect to both political and artistic speech, and freedom of religious exercise. Being able to have pleasurable experiences and to avoid non-beneficial pain.

**5. Emotions.** Being able to have attachments to things and people outside ourselves; to love those who love and care for us, to grieve at their absence; in general, to love, to grieve, to experience longing, gratitude, and justified anger. Not having one's emotional development blighted by fear and anxiety. (Supporting this capability means supporting forms of human association that can be shown to be crucial in their development.)

**6. Practical Reason.** Being able to form a conception of the good and to engage in critical reflection about the planning of one's life. (This entails protection for the liberty of conscience and religious observance.)

**7. Affiliation. A.** Being able to live with and toward others, to recognize and show concern for other human beings, to engage in various forms of social interaction; to be able to imagine the situation of another. (Protecting this capability means protecting institutions that constitute and nourish such forms of affiliation, and also protecting the freedom of assembly and political speech.)

**B.** Having the social bases of self-respect and non-humiliation; being able to be treated as a dignified being whose worth is equal to that of others. This entails

provisions of non-discrimination on the basis of race, sex, sexual orientation, ethnicity, caste, religion, and national origin.

**8. Other Species.** Being able to live with concern for and in relation to animals, plants, and the world of nature.

**9. Play.** Being able to laugh, to play, to enjoy recreational activities.

**10. Control Over One's Environment.**

**A. Political.** Being able to participate effectively in political choices that govern one's life; having the right of political participation, protection of free speech and association.

**B. Material.** Being able to hold property (both land and movable goods), and having property rights on an equal basis with others; having the right to seek employment on an equal basis with others; having the freedom from unwarranted search and seizure. In work, being able to work as a human being, exercising practical reason and entering into meaningful relationships of mutual recognition with other workers.

**REFERENCES**

- Bliss, Christopher. 1993. 'Lifestyle and the Standard of Living'. In *The Quality of Life*, pp. 415–36. M. Nussbaum and A. Sen (eds.). Clarendon Press
- Brandt, Richard B. 1979. *A Theory of the Good and the Right*. Clarendon Press
- Elster, Jon. 1982. 'Sour Grapes'. In *Utilitarianism and Beyond*, pp. 219–38. Amartya Sen and Bernard Williams (eds.). Cambridge University Press
- Elster, Jon. 1983. *Sour Grapes*. Cambridge University Press
- Friedman, Milton. 1984. 'The Methodology of Positive Economics'. In *The Philosophy of Economics*, p. 210. Daniel M. Hausman (ed.). Cambridge University Press
- Hampton, Jean. 1993. 'Feminist Contractarianism'. In *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, pp. 227–56. L. Antony and C. Witt (eds.). Westview
- Harsanyi, John C. 1982. 'Morality and the Theory of Rational Behavior'. In *Utilitarianism and Beyond*, pp. 39–62. Amartya Sen and Bernard Williams (eds.). Cambridge University Press
- Hume, David. 1739 [1888]. *A Treatise of Human Nature*, Book II, Part III, Section 3. Oxford University Press
- Mill, J. S. 1859. *On Liberty*. Indianapolis: Library of Liberal Arts
- Mill, J. S. 1869. *The Subjection of Women*. S. M. Okin (ed.). Hackett
- Nussbaum, M. 1997. 'Flawed foundations: the philosophical critique of (a certain type of) economics'. *University of Chicago Law Review*, 64:1197–14
- Nussbaum, M. 2000. *Women and Human Development: The Capabilities Approach*. Cambridge University Press
- Rose, Kalima. 1992. *Where Women Are Leaders: The SEWA Movement in India*. Zed Books
- Scanlon, T. M. 1993. 'Value, Desire and Quality of Life'. In *The Quality of Life*, pp. 201–7. M. Nussbaum and A. Sen (eds.). Clarendon Press
- Sen, Amartya. 1984. 'Rights and Capabilities'. In *Resources, Values, and Development*, pp. 307–24. Blackwell and Harvard University Press
- Sen, Amartya. 1984. 'Family and Food: Sex Bias in Poverty'. In *Resources, Values, and Development*, pp. 346–65. Blackwell and Harvard University Press
- Sen, Amartya. 1995. 'Gender Inequality and Theories of Justice'. In *Women, Culture, and Development*, pp. 259–73. M. Nussbaum and J. Glover (eds.). Clarendon Press