

# INDIVIDUAL PREFERENCES AND SOCIAL JUSTICE\*

Álvaro de Vita

## I

What is the standard by which we are to evaluate and compare an individual's quality of life? On what aspects of a person's life are we to concentrate in order to assess the level of welfare that they enjoy or, compared to what is within reach of others (in the same society, but also in other societies), they ought to enjoy?

One answer, which remains immensely influential in the welfare economics and even in contemporary political theory, is what John Harsanyi has dubbed "principle of preference autonomy": "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, 1982, p. 55). This is a disturbing answer for all those who, like me, believe that there ought to be a non-subjective basis for appraising and comparing levels of welfare for different individuals (and for different societies). It goes without saying that economic and social rights are included in the Universal Declara-

tion of Human Rights on the basis of the implicit assumption that such a basis exists.

One would be hard-pressed to demonstrate directly the plausibility of an "objective" conception of welfare. What is possible, in the first place, is to demonstrate why the subjective conception (that is to say, the conception that understands welfare solely in terms of the preferences of the agents themselves) fails to offer a plausible interpretation of individual welfare and, above all, fails to enable an acceptable basis for the comparison of levels of welfare. Secondly, we must ask ourselves whether there isn't a distinct way of dealing with the central issue (which I will set forth later) for which the subjective conception presents itself as a proposed solution. These are the two points I shall address in this essay. And although it is not my specific concern, I believe that at least some of the arguments I propose to reject the notion of individual preference as a sufficient basis for judgments of social welfare can be extended to relativist conceptions of welfare (that is, those that view their assessment as inseparable from local beliefs and moral traditions).

Two clarifications are in order with regard to the present work. The discussion I carry out below concerns some of the issues connected with the

\* Published originally in *Revista Brasileira de Ciências Sociais*, n. 29, October 1995, pp. 159-174.

Translated by Christopher Ballantines and revised by the author.

philosophy of welfare. If my argument is reasonably successful, one might obtain from it an answer to the following question: why is it more justifiable to evaluate individual welfare in terms of the access people have to certain goods, resources and opportunities (such as those laid forth in the social rights of the UN charter) than to evaluate it (solely) in terms of the satisfaction of individual preferences? The answer to this question might provide a moral basis for interpersonal comparisons of welfare. To know whether this allows for the derivation of a precise measuring-stick to evaluate public policy constitutes a separate question altogether (which I shall not attempt to address, to the possible frustration of social scientists).

As for the second clarification, I shall address herein the place that individual preferences ought to occupy in interpersonal comparisons of welfare, judgments from which we cannot shy in cases where it is necessary to assess the distribution (or to evaluate the distribution that effectively has been realized) of scarce social resources. This is a standpoint appropriate to a theory of social justice. Another point of view by which one might consider individual preferences is to see them as inputs of the political process. From this standpoint, our foremost concern lies with the institutional mechanisms for collective decision-making. Among the kinds of questions to be asked along the way are these: should the collective decision-making process be organized to aggregate and express — as faithfully as possible — the *effective* preferences of citizens; or should it be organized so as also to permit citizens to create or revise their own preferences regarding public issues? These are questions that should occupy an important place in our reflections on democracy.<sup>1</sup> Strictly speaking, however, they do not pertain to the domain of justice.<sup>2</sup> What a theory of social justice ought to facilitate is the refinement of standards by which it would be possible to judge the moral quality of the *results* of the political process.

## II

We would like to know why judgements about social justice cannot be based *merely* on

the satisfaction of individual preference; or, to state matters otherwise, what reasons are there to reject the idea that social utility can only be conceived of in terms of the aggregation of individual utilities. This is Harsanyi's (1982, p. 54) view: "The utilitarian theory I have proposed defines social utility in terms of individual utilities, and defines each person's utility function in terms of his personal preferences. Thus, in the end, social utility is defined in terms of people's personal preferences."<sup>3</sup> However, before critiquing this position, we must first determine the place from which the utilitarianism of preferences derives its plausibility.

The first point to clarify is that Harsanyi's principle has no need to base itself on a benthamite hedonist psychology. Hedonist and preferential utilitarianisms share only one point in common. Both conceive of the individual's welfare in a subjective manner: while the former envisions it in terms of the presence of certain conscious mental states, of pleasure or of pain, discernible through introspection, the latter conceives it in terms of the satisfaction or frustration of preference. (Such mental states and preferences would belong to the person whose welfare is being assessed.)

The similarities, nonetheless, end there. Unqualified hedonism offers a far too implausible explanation for individual welfare. We do a great many things independent of the subjective sensations of pleasure or pain they may produce in us. There are cases of exceptional achievement, for instance, whether professional, scientific or artistic, that can only occur after considerable personal sacrifice to their authors — either because they impose suffering or because they require the forbearance of things that produce conscious states of pleasure. This is not to mention examples of human experience which do not necessarily entail such lofty states of individual excellence. I can do things whose goal is to produce desirable states in others, or even to contribute to some objective state of things. If we feel forced to say that all of the foregoing occur *at the expense* of individual welfare we must then suspect that at least a part of the problem lies in the conception — a hedonist one, by the way — of welfare that we have adopted.

John Stuart Mill tried to escape these difficulties of the hedonist conception by discerning “higher pleasures” from “lower pleasures” and postulating that man, under normal conditions, draws greater utility from the higher pleasures and from the exercise of their higher capabilities. “It is better to be a human being dissatisfied, than a pig satisfied; better to be a Socrates dissatisfied, than a fool satisfied. And if the fool or the pig are of a different opinion, it is because they only know their own side of the question. The other party to the comparison knows both sides.” (Mill, 1961, p. 333). Through his doctrine of higher and lower pleasures, Mill in fact is adopting an objective conception of human happiness on the assumption that the ingredients for this conception are the same as those producing desirable conscious states, or those towards which individuals develop the more intense preference. Mill’s view (which we might designate as “objectivist utilitarianism”) is therefore removed from the subjective conceptions of welfare that I am confronting.

On this point we can see where the appeal for preferential utilitarianism arises. For when positioned against unqualified hedonism, the notion of utility ceases to have a necessary connection with the sensory states of the agent. An extreme example, observed by Griffin and Parfit, helps shed light on this. Wracked by pain from an advanced cancer, Freud, near the end of his life, saw himself facing a choice: to take painkillers, which would plunge him in a state of mental torpor, or not to take them and, although tormented by the pain, to retain the ability to think clearly. Freud chose the second alternative. From the standpoint of unqualified hedonism, we would say that the satisfaction of this preference diminished Freud’s level of welfare; according to preference utilitarianism, we would say that his choice was the best, all things considered, for his own welfare.

This example illustrates the strength of Harsanyi’s principle, which makes for a more frontal challenge to the “utilitarian objectivism” that I have attributed to Mill.<sup>4</sup> It is the ideal of *neutrality*: in our judgments of social justice (or “social utility”), we should avoid value judgments aimed at individual preference and choice. The better to clarify what is at issue here, let us suppose that we are in

a position to decide who gets what with regard to the distribution of a given scarce social resource. The neutrality ideal rejects the option to distribute resource  $X$  (or more of resource  $X$ ) to person  $A$ , instead of distributing it to person  $B$  (or the option to give  $B$  less of  $X$ ) because we believe that  $A$ ’s preferences are more valuable (because we believe that  $X$  will be used to satisfy preferences that we deem more valuable). Let us say that the decision at hand is whether to distribute a basket of goods to  $A$  or to  $B$  (assuming that both are in similar circumstances). Neutrality would prevent us from delivering the basket to  $A$ , and not to  $B$ , because we know — as a matter of fact — that  $B$  has a tendency to drink (and could also trade the basket for booze) and because we believe — as a value judgment — that the preference for temperance is better than the preference for drunkenness. The preference-autonomy principle requires that we find another basis for our decision. I will return to this point later.

The plausibility on face value of the utilitarianism of preferences results from the answers it provides — at first sight, in inarguable fashion — to the problem of neutrality (as it was characterized in the preceding paragraph) in comparisons of individual welfare. The power of this response lies in its rejection of paternalism (in the form of “I know what’s best for you”), especially in its connection with public policy decisions. We are not willing to concede to public authorities the discretionary powers to rule over the distribution of scarce resources — and over the use or forbearance of use of collective coercion in a general manner — on the basis of what they believe that each of us should want or do.

And so it is paternalism that constitutes the problem to which preference utilitarianism avails itself as the answer. If we wish to arrive at a non-subjective understanding of individual welfare (and of the comparison of levels of welfare) and also reject the discretionary interference in individual preferences and choices, then we cannot shrink from envisioning unique solutions to the problem of paternalism. Before we address this task, however, it is time we saw where Harsanyi’s principle fails.

### III

One point remains to be clarified in my comparison between hedonist and preference utilitarianisms. Whereas hedonism (subjective or objective) was held solely to be a conception of individual welfare — that is to say, a theory of what makes life good for whoever lives it — preference utilitarianism was viewed not only in this light, but also as a *moral* theory. To say that a thing is good for a person (that it is in their interest) because it produces in them a conscious state of pleasure fails to tell us anything at all about the kind of claim this interest presents to others.

Harsanyi's principle, however, from the outset is one of a distributive nature. What he proposes is that the satisfaction of individual preferences ought to be erected on the ultimate criteria for the distribution of scarce resources (and, generally speaking, for any and all use of collective coercion). When making public policy decisions, or when assessing them from an impartial point of view, we should only take into account the *intensity* of individual preferences and their distribution in the community and abstain — as we have seen in the foregoing section — from judging these preferences on their intrinsic merits. The basis for our distributive decisions should be constituted solely from the force and the degree of diffusion of individual preferences.

One can doubt whether it is at all possible to carry out interpersonal comparisons of welfare that take into account the intensity of individual preferences. Even if we agree that individual welfare should be assessed solely in terms of utility, and the latter solely in terms of the satisfaction of individual preferences, the problem would still remain as to which aggregate procedure we should adopt to calculate a “social function of utility”. The rule of the majority, a possible candidate to fulfill this role, is insensitive to the intensity of preferences.<sup>5</sup> However, at this point I do not wish to dwell on matters of practicality, for a principle may hold practical importance even though we lack an algorithm with which to apply it.

The strongest objections to the utilitarianism of preferences are all related to the following

question: on what basis — and without straying too far from the “principle of preference autonomy” — can we *exclude* certain preferences from our judgments of social justice? This question immediately raises another: what kinds of preferences can we reasonably exclude from these judgments? I shall examine three types of preferences that challenge any plausible conception of justice. The first category consists of *offensive* preferences; the second, of those preferences whose satisfaction imposes *excessive* claims upon others (the case of “expensive tastes”); and the third, of preferences whose satisfaction imposes claims that are more *modest* than would have been the case under different circumstances.

To comment briefly on offensive preferences, we understand them as being of a discriminatory nature and/or whose satisfaction (always bearing in mind public policy decisions) is harmful to life, to human dignity or to the freedom of others. Are we willing to admit, for instance, that attitudes such as “prisoners are better off dead” may have unimpeded currency to determine how collective coercion is to be employed? Of course not. The classical liberal strategy to deal with this sort of problem consists of protecting certain individual interests, protecting them from welfare-oriented calculations that contain an offensive component of this type by means of a constitutional charter of rights and even possibly by means of judicial review of laws approved by legislatures.<sup>6</sup> Harsanyi (1982, p. 56) confronts this problem by suggesting that human compassion, upon which utilitarian morals are founded, legitimizes the exclusion of “all clearly anti-social preferences, such as sadism, envy, resentment and malice” from the calculus of social utility.

Even if we accept this first qualification to the principle of preference autonomy, graver difficulties arise when we move to the second kind of preference mentioned above. Subjective welfarism proposes to take into account variations in individual preference. Suppose we believe that a basic objective should be to make people as equal as possible in the degree of satisfaction they achieve from their own preferences. The problem that emerges is this: on what basis are we to arbitrate

the reasonableness of demands that the satisfaction of various preferences poses before others, indeed, before society? To adapt an example suggested by Kenneth Arrow (1973, p. 254), consider a case where one individual satisfies his gastronomic preferences with water and soybean flour while another feels terribly unhappy without fine meals and fine wines. If we adopt the intensity of preferences as a standard to evaluate individual welfare and wish to equalize the level of welfare for both individuals we must then incline — however counter-intuitively — to propose that a greater portion of scarce social resources be targeted to the individual (in this case, the second) who is the more efficient consumer of resources — that is, the individual who is able to derive greater utility than the other from the consumption of the same basket of resources (Scanlon, 1975, p. 659).

It is counter-intuitive to maintain that a person who has developed “expensive tastes” ought to do justice to a greater parcel of scarce social resources in order to achieve the same level of satisfaction as that enjoyed by a person of more modest tastes and ambitions. Nevertheless, it is not easy to see how subjective welfarism could stand up to this objection.<sup>7</sup> As Dworkin (1981, pp. 185-246) forcefully argued, conceptions of welfare equality — conceptions, that is to say, that propose to equalize individual welfare according to the subjective appraisal that each individual makes of his own level of welfare — are in the end contradictory. In interpersonal comparisons of welfare, whatever the conception of equality we adopt, we are led inevitably to evaluate the point up to which the social claims arising from the satisfaction of certain preferences are reasonable. What is reasonable, however, cannot be discerned on the basis of the attributes of the preferences nor of their intensity — *it is a pattern external to the preferences themselves*. If we cannot avoid falling back to a notion of “reasonable claims” in interpersonal comparisons, then our concept of equality, although held to a subjective standard, already contains prior built-in assumptions about distributive equity. And these prior assumptions require justification independent of the individual interests to be compared.

Harsanyi does not ignore the objection of expensive tastes: “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.” But he faces this objection by invoking a distinction that cannot easily be squared with his own principle of preference autonomy:

[...] All we have to do is to distinguish between a person's manifest preferences and his true preferences. His manifest preferences are his actual preferences as manifested by his observed behavior, including preferences possibly based on erroneous factual beliefs, or on careless logical analysis, or on strong emotions that at the moment greatly hinder 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. (Harsanyi, 1982, p. 55)

It is hard to imagine that this distinction between manifest preferences and true preferences does not eventually constitute the thin end of the wedge precisely for that which the principle of preference autonomy above all rejects: paternalism. The temptation faced by the utilitarian policymaker would be to start from his preferred social utility function and then to postulate that it is in accord with the individual utilities that derive from the satisfaction of the “right” preferences. The point Dworkin emphasizes stands: when we compare individual interests, the standard we employ is not contained within those interests. And it would do very little good to replace the standard of satisfaction of individual preferences with the standard of preferences held by those who make policy decisions.

The third kind of preference I have laid forth above (and which might be grasped as an important subset of the second kind) remains to be considered. How can we evaluate preferences

whose satisfaction requires less from others if institutional background circumstances were otherwise? This is the problem that Elster (1983, pp. 109-140) has called “sour grapes”. A person may attach little value to a good, a right or an opportunity; they may not even appear in his scale of preferences for the simple reason that he cannot easily develop a preference for something he cannot see — fairly realistically — as part of his circumstances in life. It is to ask for too much that somebody would have a preference for alternatives they cannot see included within their available options.

This applies above all when we attempt to evaluate the preferences of people in vulnerable positions. Consider, for example, the case of a boy in Northeastern Brazil who, because he must work, cannot attend school. In his individual scale of utility, and in that of his parents, it is quite possible that “increasing income” ranks more highly than “attending school”. If we adopt a welfare-oriented standard to evaluate welfare, we would be induced to believe that it would be best to do whatever possible to increase the boy’s income. Instead, however, we might raise the question: “what changes in the circumstances of the boy’s life are necessary so that education will figure as an important value in his (and his parents’) scale of preferences?” I cannot see how this question might be raised from the standpoint of the utilitarianism of preferences.

The objection we are raising to the welfarist perspective is this: the satisfaction of individual preferences is not an adequate guide for public policy decisions because what people prefer is itself the result, by and large, of the goods, resources and rights with which they already have been provided through public action. Cass Sustein (1991, p. 8) remarks that there is solid empirical evidence in economics and social psychology for what can be termed “the endowment effect”, that is, the existence of a causal link between individual attitudes regarding certain goods and rights and the supply of the same goods and rights. “The endowment effect is the consequence, for preferences and willingness to pay, of the initial allocation of an entitlement.”

The relevance of the endowment effect to the evaluation of the form in which individual preferences should count in policy decisions is great. We could point to countless examples that refer to distinct areas of state intervention. The concession of labor rights to domestic workers strongly influences the perception they have of their own activity and of what their employers can demand of them; the extension of pensions to rural workers affects their preferences regarding the value of the pensions system; attitudes of slum-dwellers toward the environment are heavily influenced by the provision (or non-provision) of what, in their view, are the essential public goods, such as potable water, waste-removal and public sanitation; the perception by residents in peripheral districts of São Paulo of the importance of access to medical care can be affected by the supply of health services in those districts;<sup>8</sup> and so on.

Let us admit that it is correct to argue that the supply of certain goods, resources, rights, legal norms or opportunities broadly conditions the attitudes individuals have in relation to each of these things. In this case, it is simply false to state that the distribution of these same goods, resources, rights, legal norms or opportunities can be justified by the satisfaction of individual preferences. *The equity of a particular distribution cannot be justified by the preferences that individuals are induced to cultivate for this same distribution.* Once again, through the same mechanism we have already witnessed in the exercise of “expensive tastes”, the problem posed by the malleability of preferences pressures us to find a non-subjective basis for the assessment and comparison of levels of individual welfare. Before we tackle this, however, there is a final point to be elucidated.

#### IV

The malleability of preferences raises objections not only for subjective welfarism (which places a full emphasis on interpersonal variations of preference). A similar objection applies to relativist conceptions of social welfare (which place full emphasis on the intercultural variation of moral standards).

I have argued above that we have considerable reasons to reject a principle that advises for the satisfaction of individual preferences without taking into account the fact that they are conditioned by socio-economic background and by the mode of state intervention (or its absence). Similarly, there are powerful reasons to reject conceptions that propose that a person's welfare can only be evaluated according to the values and beliefs of the community to which he belongs, especially if: (1) these values and beliefs perpetuate one's situation of inferiority and oppression within this community; and (2) one's life circumstances hardly permit any option other than adherence to these values and, consequently, impart a positive value to one's own oppression. Moreover, in this case, the fact that a person "prefers" the situation in which he is already placed is not a good reason — and even less a just reason — for guiding our judgment about what he is entitled to on the basis of his "preference".

To illustrate this point, consider the following example. The dominance of certain moral traditions and certain conceptions of family life leads to the existence in many a poor country not only of an unjust social distribution of resources but — above and beyond this — an unjust *intra-family* distribution of resources. Among poor families, the distribution of food, medical attention, and educational opportunity generally benefits men and income-earning adults, firstly, as it privileges boys to the detriment of girls secondly (World Bank, 1990, p. 37). Suppose (quite plausibly) that the community beliefs and moral traditions that legitimize this state of affairs are beyond dispute, to the point where women themselves see no injustice in it. With this scenario in mind, how could policy decisions base themselves in the assessment these women make of their own situation? Would we be willing to assert, for example, that to provide them with educational opportunities is not a pressing objective (or not as pressing as it is for men) because their own way of life fails to valorize women's education? This is a clear example, in my view, of how the adoption of individual attitudes — conditioned, in this case, by adherence to local custom — as the ultimate criterion for judgments of social welfare can lead to absurd distortions.

Let me add, in passing, that this kind of error in judgment, which easily could be committed by relativists, does not occur in the 1990 *World Development Report*. The report recommends that, in poor countries, the state invest directly in the improvement of quality of life for poor women (who are among the greatest victims of relativist conceptions of social justice). This recommendation is consistent with the considerations pertaining to social justice that I mentioned in the preceding paragraph. It is grounded, moreover, in a fact of the utmost importance. Comparing the way in which poor men and women spend their income, it is observable that a greater portion of women's income is expended on the family's welfare. The report's conclusion is clear: to increase women's income directly not only enhances their bargaining power within the family but also provides a sound means for benefiting their children (World Bank, 1990, p. 37).<sup>9</sup>

There is something emphatically wrong with conceptions of social justice that weaken the position of those who are most vulnerable. This, in short, is the principal objection that should be raised before utilitarianism of preferences and relativist conceptions of social welfare. "Any relativism", as Onora O'Neill (1993, p. 304) puts it, "tends to prejudicate the position of the weak, whose weakness is mirrored and partly constituted by their marginalization in received ways of thought and by their subordination and oppression in established orders." What indeed remains beyond comprehension is the fascination which relativism seems to exercise over certain varieties of left-wing political thought.<sup>10</sup>

## V

So far I have argued that the two subjective theoretical perspectives for the evaluation and comparison of welfare levels — hedonism and the utilitarianism of preferences — are inherently contradictory. When we ask ourselves (always from the standpoint of decision-making in public policy) which conscious states ought to be produced in individuals, or which individual preferences ought to be satisfied, we are compelled to use — even if

only implicitly — a non-subjective standard: that is, some standard by which we can judge the reasonableness, in terms of the claims it poses on others, of the production of pleasurable states of consciousness and of the satisfaction of individual preferences.

If we abandon the subjective standard, we enter into a field of conceptions of welfare that are affiliated to what Derek Parfit has called “objective list theory”. According to this view, there are certain things that are good or bad for human life, aside from people’s desire to pursue the good and to avoid the bad. “The good things might include moral goodness, rational activity, the development of one’s abilities, having children and being a good parent, knowledge and awareness of true beauty. The bad things might include being betrayed, manipulated, slandered, deceived, being deprived of liberty or dignity, and enjoying either sadistic pleasure, or aesthetic pleasure in what is in fact ugly.” (Parfit, 1984).

What we are considering, then, is a theory of individual welfare rather than one — just yet — of morality. Nothing has been said, so far, about the duties these ingredients for a good life impose on others. (To be a good father or a good mother, for instance, however great their importance in individual welfare, is not a good with which others can provide us.<sup>11</sup>) Wouldn’t such a theory — and this is the objection to consider from the outset — itself be based in an unabashed value judgment as to what comprises a good life?

It is true that utilitarianism of preferences embraces the assumption that value judgments — if this perspective is adopted — are to be avoided whenever possible. For Harsanyi, the identification and comparison of the intensity of human preferences are judgments of fact.<sup>12</sup> However, to adopt objective list theory does not mean that we must henceforth surrender to a value judgment. In this regard, it is pertinent to recall Amartya Sen’s remark to Mollie Orshansky that “poverty, like beauty, lies in the eye of the beholder”. What Orshansky means is that our evaluation of poverty as something bad (and that certainly occupies a prominent place in the list of bad things mentioned above) itself constitutes a value judgment. But on this point there is

some confusion surrounding two types of judgment. A direct and prescriptive judgment — “this is bad” — is one thing. But a judgment along the lines of “according to broadly held normative standards, this is bad” is quite a different thing indeed. Although our values always interfere, one way or another, with the manner in which we evaluate the facts, we should bear in mind that “to describe a prevailing prescription is an act of description, not prescription” (Sen, 1981, p. 17). If subjective welfarism claims to be grounded in facts “as they are” (see, however, note 12), objective list theory seeks its support through *facts about values*.

It is no sinful excess of objectivity to maintain that certain things are valuable and others harmful to a variety of individual conceptions of the good (though not all of them) and to human life in a variety of cultural contexts (though not all of them). Assumptions of this sort are present in all conceptions of welfare belonging to the family of the objective list theory. Among its most prominent members are: Rawls’ notion that individual welfare is to be assessed by reference to an index of “primary goods”;<sup>13</sup> Sen’s proposition that the focus be directed toward “a set of human functionings and capabilities”; and Doyal and Gough’s conception (and others similar to it), which proposes to evaluate welfare in connection with certain basic human needs.<sup>14</sup> All of these conceptions are non-subjective and anti-relativist; and all of them can be understood as interpretations of the conception of human welfare that lies at the foundation of the Universal Declaration of Human Rights.

And so how are we to compare conflicting individual interests from the standpoint of objective list theory? When we raise this question, our conception of individual welfare becomes the keystone of a moral theory (a theory of distributive equality). The fundamental point to be stressed is this: unlike what subjective welfarism proposes, we do not judge the legitimacy of the claims which these interests pose regarding scarce social resources according to the force with which their parties defend it, nor by the intensity with which they are preferred, nor even by the degree of satisfaction that the fulfillment of these claims might bring about. Our task is to inquire into the reasons for



which something is preferred (Scanlon, 1975, 1991 and 1993). If we wish to escape the dead-end into which subjective standards lead us, there is no way to avoid judgments regarding the content of preferences in interpersonal comparisons. That someone (or some group) prefers something is not a good reason, in and of itself, for this interest to weigh in public policy decisions. We must still establish a judgment of the moral importance of the interest in question. And we do so by recourse to our conception of individual welfare (“ours” in the sense of those who accept objective list theory). An interest will entail greater moral weight on the basis of its greater relationship to one or several goods that are perceived, by people with distinct values, as ingredients indispensable to a good human life. According to this view, it is not merely by being the object of preference that a thing is good or valuable. Because it constitutes a good, we have a reason — intersubjective in nature — to prefer it.

Note that to judge the moral importance of individual interests in interpersonal comparisons does not involve judgment of the intrinsic value of these interests, nor determining the point at which the interested parties are able correctly to evaluate what is best for them. Judgments of the moral urgency of an individual preference or interest are perfectly compatible with the benthamite formula that “the individual is the best judge of his own good”. If a person feels deeply unhappy for not being able to enjoy rare and sophisticated food and drink, we would not say that he or she has a preference for “inferior” pleasures. Nor would we assert that such a preference reveals that he does not know what is best for him. The only thing we do judge are the claims that the satisfaction of this preference makes on scarce social resources. When we inquire, for the purposes of interpersonal comparison, into the reasons why something is intensely desired, we are to bear in mind solely the social opportunity costs of satisfying this desire — a cost which is to be evaluated by reference to the desires, intense or otherwise, that will be frustrated. (I will return to this point in the following section.)

Leaving aside purely strategic considerations,<sup>15</sup> objective list theory provides for a more plausible explanation of the basis of comparisons of

individual interest that we frequently conduct. Normally we consider that preserving freedom of expression is a more important good than satisfying the preferences (however intense they may be) of people who hate to see nudity or sex on television; that guaranteeing everyone sufficient opportunity to develop one’s basic capabilities is more pressing than satisfying the preferences of some for especially costly forms of education or training; and that it is more important to assure proper nutrition for the hungry than it is to guarantee others the means to satisfy a given preference derived from religious belief (such as the construction of a temple). In all of these cases, we are implicitly resorting to a non-subjective standard for evaluating individual interests, according to which some of these interests will be seen as private preferences (i.e., that they do not pose claims before society) while others will be recognized as rights (i.e., as aspects of individual welfare that impose responsibilities on others).

## VI

Suppose we agree on the existence of a non-subjective moral foundation, such as that proposed by objective list theory, for comparisons of interpersonal welfare. One question remains to be answered. Is it possible, by resorting to this foundation, to envision another way of addressing the problem that the utilitarianism of preferences sets out to resolve: namely, the problem of neutrality (as it was characterized in section II)? Would it be unacceptably paternalistic to propose that the evaluation of a person’s standard of living be conducted not according to his own subjective assessment of his situation but rather by the access assured to him to certain goods, resources and opportunities — especially including, among others, the social rights recognized in the UN charter — that we believe to be ingredients of a good life. Might we thus be issuing a value judgment — that “what So-and-so wants is not in his best interest” — about individual preferences and attitudes? What role do we assign, after all, to individual responsibility in a person’s life?

I shall sketch out the general line of argument adopted by all conceptions of individual welfare,

to my knowledge, that accept objective list theory. For this purpose, I will start with an inconsistency that G.A. Cohen purports to have detected in Rawls' political philosophy. According to Cohen (1993, pp. 13-14), the problem confronting egalitarian and left thought, generally speaking, "is that the picture of the individual as responsibly guiding his own taste formation is hard to reconcile with claims Rawls uses elsewhere in a fundamental way to support his egalitarianism."

Let us better clarify the two things that seem irreconcilable. On the one hand are the assumptions necessary for justifying egalitarianism. Rawls (and left thought, broadly speaking) rejects the notion that the unequal distribution of resources and opportunities can be justified on the basis of individual merit. The argument is that personal merit — an attribute inseparable from highly personal preferences, tastes and attitudes<sup>16</sup> — is intensely conditioned by certain circumstances in an individual's life that lie beyond the realm of choice. One cannot choose the country, region or community where one is born and raised, any more than one's starting-position in society, family, or genetic baggage. And so to attribute a person's unfavorable situation to his own preference (to see this as his own shortcoming), when a more careful examination of the situation would recommend attributing it largely to circumstances outside one's choice, amounts to a familiar and unacceptable instance of victim blaming. This position is a variation on the argument regarding the malleability of preferences that I used earlier to criticize subjective welfarism.

On the other hand, there is the assumption that the individual should be responsible for cultivating his own tastes, preferences and ends. Rawls resorts to this assumption to explain why our conception of distributive equality<sup>17</sup> need not preoccupy itself with the satisfaction of expensive tastes. It is worth recalling Rawls' passage on this point:

[...] as moral persons citizens have some part in forming and cultivating their final ends and preferences. It is not by itself an objection to the use of primary goods that it does not accommodate those with expensive tastes. One must argue in addition

that it is unreasonable, if not unjust, to hold such persons responsible for their preferences and to require them to make out as best they can. But to argue this seems to presuppose that citizens' preferences are beyond their control as propensities or cravings which simply happen. Citizens seem to be regarded as passive carriers of desires. The use of primary goods, however, relies on a capacity to assume responsibility for our ends. (Rawls, 1982, pp. 168-169).

For Cohen, these two components of Rawls' theory fail to mesh. Individual preferences are seen in one manner when the task at hand is to disqualify merit as an egalitarian distributive principle for the basic structure of society; they are seen in a different light when the issue is to justify the non-satisfaction of expensive tastes. Egalitarianism thus appears to run into the unpleasant failure to reconcile the assumptions underlying principles of distributive equality with those that underlie principles of individual autonomy. Indeed, one can assume that the justification for egalitarian policies would always depend on judgments that oftentimes are not easily formed regarding the degree of autonomy/heteronomy — or of individual responsibility/non-responsibility — in the formation of individual preferences. The greater the heteronomy of the individual's preferences that contribute to maintaining him in an unfavorable position,<sup>18</sup> the more authoritatively can egalitarian policies afford to ignore them. The greater their autonomy, the less can public policy decisions afford to fail to take them into account.

Why is this problem relevant to the question of neutrality? If our conception of distributive equality depends on judgments of the degree of autonomy of individual preference, we must then exercise value judgments about these preferences. Preferences generated in autonomous circumstances are "good" (from the standpoint of decisions on public policy), while those generated under non-autonomous circumstances are "bad". And to establish distinctions on the intrinsic value of individual interests and attitudes is precisely what the ideal of neutrality rejects.

There is, nevertheless, a serious misunderstanding in Cohen's objection to Rawls. We have

seen that one of the reasons for rejecting subjective welfarism derives from the problem of the malleability — or the heteronomy — of preference. This is an objection that is raised to a theory that proposes that our judgments of social justice should be governed by the degree of satisfaction (realized or yet to be realized) of individual preferences. However, along the same lines that I argued in the preceding section on the inquiry into the reasons why something is desired, interpersonal comparisons of welfare grounded in the roster of primary goods set forth by Rawls, or in other conceptions that adopt objective list theory, do not require any sort of judgment pertaining to the degree of autonomy/heteronomy of individual interests, ambitions or ends. This is not to say that individuals can only be compensated for their preferences, especially preferences that leave them in a disadvantageous position, in cases where it can be demonstrated that these preferences are due to factors or attributes outside the realm of individual choice. No judgment whatsoever is being emitted about what people deserve. What we are asserting is that the basic distribution of resources and opportunities in society should be independent of preference, *whether the latter is constituted in heteronomous or autonomous form*.

Neither judgments about the degree of injustice in the societies in which we live, nor judgments about the social well-being that *could* be constituted if we lived in a just society require a basis in evaluations of the degree of autonomy of their attendant individual preferences and conceptions of what is good. The object of justice, as Rawls remarked, is the basic structure of society and not the constitution of moral adjudications of particular cases.<sup>19</sup> What we assume is that there are goods, resources and opportunities that persons, however greatly they may diverge in their individually held attitudes and values, have reason to desire. The next assumption — adopted by all conceptions of welfare related to objective list theory — is that society's basic institutions should be organized and should function in such a manner as to assure that an equitable portion of these goods and opportunities be assured to each person. Whatever each individual makes of the opportunities thus offered

to them — the preferences he will seek to cultivate, the goals he will seek to achieve — is no longer, as a rule, any of society's business.<sup>20</sup>

Assume that an equal measure of resources and opportunities was afforded to persons *A* and *B*. Although *A* is dissatisfied with his situation because certain preferences have been frustrated, and *B* looks upon his situation positively, we would still be compelled to assert that, in spite of their divergent subjective views, *A* and *B* enjoy equal levels of welfare. Or yet, under the same conditions of distributive equality, that *A* invests all his effort in a costly career that promises financial reward or public prestige, while *B* prefers a life less driven by professional realization and more dedicated to family and friends. From a public standpoint, there is no reason to judge that *A*'s choices and preferences are more valuable than *B*'s. According to Rawls, in the passage we quoted above, given an equitable distribution of primary goods or — as Sen would have it — given institutional conditions that promote equally the development of each individual's capabilities, we may assume that individuals are capable of taking responsibility for their own objectives.

The conclusion at which we have arrived hardly supports the contradiction pointed out by Cohen (namely, that between the assumptions necessary to justify distributive equality and the assumptions of individual autonomy). *It is solely the assurance of distributive equality in the basic structure of society that enables us to avoid judgments of the intrinsic value of individual preferences and choices*. This is the stock answer that non-subjective conceptions of individual welfare can offer to problems arising from neutrality or paternalism in interpersonal comparisons of welfare — problems for which the utilitarianism of preferences, at first sight, would appear to offer the most plausible solution. Despite the fact that this formulation entails a high degree of abstraction, this response carries implications that are quite definite in nature. If we wish to avoid paternalistic stances in judgments of social justice, we had better *not* focus directly on individual preferences, attitudes and interests, and direct our attention instead to institutional conditions for the allocation of re-

sources and opportunities of value to a variety of individual conceptions of what is good. In my view, this would be the best philosophical approach to the social and economic rights contained in the Universal Declaration and in other UN documents.

## VII

One last point remains to be discussed. Up to this point I have sought to elucidate what people are entitled to under the auspices of a non-subjective standard for assessing welfare. However, the concept of neutrality and anti-paternalism assumed by this kind of standard can also be laid forth through the perspective of the *responsibilities* posed by the rights to social welfare. Consider two opposite, if not extreme, points of view on the correlation between rights and responsibilities. At one extreme we find the “libertarian” view, according to which we are under no obligation, either in our individual conduct or in the collective decisions in which we participate, to recognize any *negative responsibility* for the unfavorable circumstances of other people’s lives.<sup>21</sup> From this standpoint, we are not responsible for the damages or privations that others might suffer because we have declined to do what was within our grasp to avoid the damages or to diminish the suffering. If we ourselves do not directly cause other people’s privations — if we are not positively responsible for them — we can then ignore them and nothing can justify society’s interference in our preferences and choices. At the other extreme, there is the view that there are no limits to the negative responsibility that consideration for the life and the welfare of others might pose to us. We are responsible for any privations others might suffer (on an ever broader, planetary scale) whether we have caused them directly or not, if ever we failed to do what we could to diminish it. According to the latter view, there is no interference in personal conduct, practically speaking, that cannot be justified on the grounds of an impartial compassion that we owe to other people’s lives and welfare.

The conception of neutrality and anti-paternalism that I have set forth above combines a part

of each of these views so as to give rise to a third. It shares the libertarian preoccupation with setting limits to the claims that others establish before personal preferences and choices. I suppose, nevertheless, that many among us do not consider it good to live in a world where even the satisfaction of one of our more mundane preferences — to eat a decent meal in a moderately expensive restaurant in São Paulo — raises the issue of immorality, for our consumption of resources at one place and in one evening that would otherwise be sufficient to sustain an entire family of starving slum dwellers for a whole month. We are thus compelled to weigh the considerations that provide ballast for the second point of view.

In our personal lives, we would like to do whatever we want, without this being judged (by others or by ourselves) as a show of indifference toward others. The only possible solution consists of transferring the burden of negative responsibility, of which we would rather be free in our personal life, to society’s basic institutions. This is the third way between the two aforementioned extremes. Thomas Pogge (1995) refers to it as the “institutional focus” for the correlation between rights and responsibilities. Collective recognition of negative responsibility is the precondition for ignoring it in personal conduct. Collectively, we are responsible for what could be done to avoid the damage and privation that many of us must endure. This establishes the responsibility for supporting institutions whose goal is to assure equal levels for everyone’s welfare, assessed according to the non-subjective standard proposed above.

Where such institutions do not exist, our foremost duty is to create them, according to the institutional focus. The point about collective negative responsibility is that the status quo ceases to be seen as a benchmark to judge the legitimacy of collective decisions and public policy. The critical trick about libertarian morality consists of assuming as given the distribution of social benefits and entitlements, and to place under suspicious gaze only those injustices for which the state might be held accountable (the indiscriminate massacre of “suspects” by the police in Brazil’s major cities, for instance). The institutional focus on the correlation

between rights and responsibilities, meanwhile, proposes that we are collectively responsible for what the public authorities do in our name *and* what the common institutions under which we live are not capable of preventing (that many should suffer endemic malnutrition, or that the route to the development of their individual potential should remain blocked, for example). It is not hard to see that the establishment of collective responsibility for *what has not been done* is essential to the admission of economic and social rights — and not solely of civil and political rights — as genuine human rights.

Once negative collective responsibility is sufficiently recognized by the institutions under which we live, so the argument goes, we can then reclaim the right to live our personal lives according to the libertarian ethos of non-interference. And so we must strive to realize our preferences and choices without allowing that whosoever should question their intrinsic value. Thomas Nagel (1991, p. 84) expresses it thus: “an acceptable moral framework for apportioning negative interpersonal responsibilities is a condition of the moral acceptability of strict limitations on negative responsibility in the rules of individual conduct that govern personal relations within it.”

The line of reasoning above reaffirms the conclusion of the preceding section. We set out from the idea that our judgments of social justice should embody an anti-paternalist concern. Earlier, we saw that the best way to satisfy this preoccupation was not to focus directly on the satisfaction of individual preference (which is what subjective welfarism proposes). And now we have seen that it is not resolved by encapsulating individuals within their own preferences and choices, denying them the existence of responsibilities of other people’s welfare (which is what libertarianism proposes). In the face of these two perspectives, I have argued that there is a non-subjective basis for interpersonal comparisons of welfare and that the institutional recognition of this groundwork affords a more appropriate interpretation for our concerns about neutrality and our intent to reject paternalism. To state matters more forcefully: it is quite possible that a more equal distribution not

only of those resources and opportunities that are of value to different individual planes of being, to different ways of existence, but also of their corresponding negative interpersonal responsibilities, might be the *only* way of really taking the rejection of paternalism seriously.

## NOTES

- 1 See, for example, Elster (1983, pp. 33-42) for a discussion of these issues.
- 2 The question as to which democratic institutional devices are most likely to generate just results is, of course, of the utmost importance. In this connection, for example, see Van Parijs (1995).
- 3 This viewpoint is known, in the literature, as “subjective welfarism”.
- 4 Harsanyi’s principle generally rejects all “perfectionist” moral theories, i.e., those that posit that individual welfare should not be evaluated according to the satisfaction of agents’ effective preferences but rather on the basis of the satisfaction of those preferences conducive to the ideal of a better human life.
- 5 In the language of welfare economics, the rule of the majority is a decision-making procedure that allows for “ordinal” but not “cardinal” comparisons of utility. In the first case, it is only possible to state that alternative *a* is preferred over *b*, which in turn is preferred over *c*, and so on. In the second case, the order will take into account how *strongly* each of the alternatives is preferred. The ordering of preferences thus would take on the following form: alternative *a*, yielding 40 units of utility, is preferred over *b*, which yields 35 units of utility, which is preferred over *c*, with 5 units of utility, and so on.
- 6 Along these lines, Dworkin (1984) famously proposed to define individual rights as “trumps” which individuals can play against discriminatory utilitarian calculations; moreover, it is worth recalling Samuel Freeman’s (1994) defense — relevant not only for the US context — of the institution of judicial review
- 7 This is one of the main objections Rawls raises against utilitarianism of preferences.
- 8 This particular example is empirically supported by research on popular health movements carried out by Cedec (and coordinated by Amelia Cohn) in the 1980s among residents of the eastern and southern districts of São Paulo. In spite of the greater lack of health care facilities in the southern district, it was the residents of the eastern district who developed a more acute perception of health care as priority problem and who demonstrated greater willingness to participate in collective actions aimed at redressing healthcare rights, *precisely because there was greater public investment in health services in that area* (Cedec, 1989).

- 9 The example discussed in the two preceding paragraphs is especially pertinent to a widening debate in Brazil on poverty-reduction policy. In this connection, there seems to be a growing perception, in comparison to policies more intensely subject to clientelistic intermediation (such as the distribution of food and the creation of subsidized employment) of the advantages of cash benefits. However, the considerations regarding intra-familial distribution, which I have discussed above and which are applicable to the Brazilian case, would lead us to prefer *individual* benefits (such as the negative income tax proposed by senator Eduardo Suplicy) to *family* benefits. Individual benefits are the most direct and certain means of increasing poor women's income.
- 10 It would be worth examining, for instance, the contortions Michael Walzer (1983) must perform in order to reconcile the relativist approach with the particular aims of the social critique he is intent on pursuing.
- 11 Others, nevertheless, may provide us sufficient *opportunities*, that we may become a good parent.
- 12 Donald Davidson emphatically challenges this assumption. For him, the attribution of interests and preferences to others *already* involves an appeal to a common, shared ground: "[...] the propositions I must use to interpret the attitudes of another are defined by the roles they play in my thought and feelings and behavior; therefore in interpretation they must play appropriately similar roles. It is a consequence of this fact that correct interpretation makes interpreter and interpreted share many strategically important beliefs and values." (Davidson, 1986, p. 209).
- 13 Such as income and wealth, and opportunities in education, occupation, enjoyment of leisure and self-respect.
- 14 Rawls, 1971, pp. 90-95 and ch. 7, and 1982; Sen, 1985, 1993a and 1993b; Doyal and Gough, 1994.
- 15 The theory of justice does not take into account strategic considerations. Situations propitious to just deliberation, such as Rawls' original position, are conceived precisely to neutralize inequalities in power. They are not to interfere with the formulation of principles. It is a separate issue to know how the application of a principle of justice will affect these inequalities. This is valid as well, of course, for the utilitarianism of preferences, which would view as morally incorrect, for example, the decision to satisfy a weak preference of a small, powerful group to the detriment of a strong preference held by a more numerous group of people lacking in the resources of power.
- 16 For example: individual attitudes concerning work vs. non-work; immediate consumption vs. savings; the willingness to accept risks.
- 17 "Ours" in the sense of those accepting objective list theory. Rawls' conception, which proposes the distribution of a set of primary goods according to the "principle of difference", is itself but one among several members of larger family.
- 18 Consider the case of a poor person who prefers the realization of transport policies that mainly benefit the non-poor over policies for expansion and improvement of public services in education, health and mass transit that would benefit the underprivileged.
- 19 The constitution of judgments concerning particular cases — i.e., what a particular individual is entitled to or liable for — is an object of *legal* justice.
- 20 "As a rule", because there are those cases of justifiable paternalist interference in individual choice, such as the subjection to appropriate treatment of an individual undergoing a psychotic episode, even though this action may be contrary, at the time, to their volition. But even this involves the judgment of particular cases.
- 21 For the notion of negative responsibility, see Nagel (1991, pp. 83-84). The finest philosophical articulation of the libertarian perspective continues to be that of Nozick (1974).

## REFERENCES

- ARROW, Kenneth. (1973), "Some Ordinalist-Utilitarian Notes on Rawls' Theory of Justice". *Journal of Philosophy*, 70, 9: 245-263.
- CEDEC. (1989), *Informações Cedec*, 1.
- COHEN, Gerald. (1993), "Equality of What? On Welfare, Goods and Capabilities", in Amartya Sen and Martha Nussbaum (eds.), *The Quality of Life*, Oxford, Clarendon Press.
- DAVIDSON, Donald. (1986), "Judging Interpersonal Interests", in Jon Elster and Aanund Hylland, *Foundations of Social Choice Theory*, Cambridge, Cambridge University Press.
- DOYAL, Len and GOUGH, Ian. (1994), "O Direito à Satisfação das Necessidades". *Lua Nova*, 33.
- DWORKIN, Ronald. (1981), "What Is Equality?" (Part 1: "Equality of Welfare"). *Philosophy and Public Affairs*, 10, 3.
- \_\_\_\_\_. (1984), "Rights as Trumps", in Jeremy Waldron (ed.), *Theory of Rights*, Oxford, Oxford University Press.
- ELSTER, Jon. (1983), *Sour Grapes*. Cambridge, Cambridge University Press.
- FREEMAN, Samuel. (1994), "Democracia e Controle Jurídico da Constitucionalidade". *Lua Nova*, 32.
- HARSANYI, John. (1982), "Morality and the Theory of Rational Behaviour", in A. Sen and Bernard Williams (eds.), *Utilitarianism and Beyond*, Cambridge, Cambridge University Press.

- MILL, John Stuart. (1961), "Utilitarianism", in Marshall Cohen (ed.), *The Philosophy of John Stuart Mill*, New York, The Modern Library.
- NAGEL, Thomas. (1991), *Equality and Partiality*. Oxford, Oxford University Press.
- NOZICK, Robert. (1974), *Anarchy, State and Utopia*. New York, Basic Books.
- O'NEILL, Onora. (1993), "Justice, Gender, and International Boundaries", in A. Sen and M. Nussbaum (eds.), *The Quality of Life*, Oxford, Clarendon Press.
- PARFIT, Derek. (1984), *Reasons and Persons*. Oxford, Clarendon Press.
- POGGE, Thomas. (1994), How Should Human Rights Be Conceived? Manuscript.
- RAWLS, John. (1971), *A Theory of Justice*. Cambridge, Harvard University Press.
- \_\_\_\_\_. (1982), "Social Unity and Primary Goods", in A. Sen and B. Williams, *Utilitarianism and Beyond*, Cambridge, Cambridge University Press.
- SCANLON, Thomas. (1975), "Preferences and Urgence". *Journal of Philosophy*, 72: 655-669.
- \_\_\_\_\_. (1991), "The Moral Basis of Interpersonal Comparisons", in Jon Elster and John Roemer (eds.), *Interpersonal Comparisons of Well-Being*, Cambridge, Cambridge University Press.
- \_\_\_\_\_. (1993), "Value, Desire and Quality of life", in A. Sen and M. Nussbaum (eds.), *The Quality of Life*, Oxford, Clarendon Press.
- SEN, Amartya. (1981), *Poverty and Amine*. Oxford, Clarendon Press.
- \_\_\_\_\_. (1985), "Well-Being, Agency, and Freedom". *The Journal of Philosophy*, 82.
- \_\_\_\_\_. (1993a), "O Desenvolvimento como Expansão de Capacidades". *Lua Nova*, 28-29.
- \_\_\_\_\_. (1993b), "Capability and Well-Being", in A. Sen and M. Nussbaum (eds.), *The Quality of Life*, Oxford, Clarendon Press.
- SEN, Amartya and NUSSBAUM, Martha (eds.). (1993), *The Quality of Life*. Oxford, Clarendon Press.
- SEN, Amartya and WILLIAMS, Bernard (eds.). (1982), *Utilitarianism and Beyond*. Cambridge, Cambridge University Press.
- SUSTEIN, Cass. (1991), "Preferences and Politics". *Philosophy and Public Affairs*, 20, 1: 3-34.
- VAN PARIJS. (1995), "A Democracia e a Justiça São Incompatíveis?". *Revista de Estudos Avançados*, USP, 23.
- VITA, Álvaro de. (1993), *Justiça Liberal. Argumentos Liberais contra o Neoliberalismo*. São Paulo, Paz e Terra.
- WALZER, Michael. (1983), *Spheres of Justice*. Oxford, Blackwell.
- WORLD BANK. (1990), *World Development Report*.