

Printed in *Competition and Cooperation: Conversations with Nobelists about Economics and Political Science*. eds James Alt, Margaret Levi, and Elinor Ostrom. New York: Russell Sage Foundation. pp. 4 - 33.

***Kenneth Arrow, Welfare Aggregation and Progress in Political Theory***<sup>1,2</sup>

by

Norman Frohlich and Joe Oppenheimer

Faculty of Management, University of Manitoba

Department of Government & Politics, University of Maryland

**ABSTRACT**

Kenneth Arrow's Social Choice and Individual Values and his related writings are best known in Political Science for demonstrating the impossibility of constructing an ideal democratic procedure for aggregating individual choices. A broad concern for the relationship between individual values, individual limitations and the quality of social life underlies his work. Arrow builds on the notion that to tie social welfare to democratic political decisions one must both be concerned with questions of aggregation, and also have a firm understanding of individual choice. Hence, his careful delineation of the meaning of preferences, his identification of desiderata for social welfare functions, and his subsequent inclusion of principles of distributive justice in his analysis of social welfare functions (or constitutions, as he was to come to call them). Thus Arrow's work on social choice and social welfare intersects the agenda of modern political science and traditional political philosophy at a number of crucial junctures.

Recent work in related fields raises questions regarding the uniqueness and stability of preferences and hence requires a substantial reformulation of the theoretical underpinnings of the field. Specifically, research in cognitive psychology indicates that preferences over outcomes is determined by framing, or cues. Expressions of preferences over a set of alternatives can only be understood when relativized to various sets of cues. The epistemological and normative interpretation of such revealed preferences, and in particular the question of which preferences are to be counted and aggregated becomes problematic. Arrow's discussion of individual preferences and the existence of a 'universal set': a homogeneous preference structure for all individuals, furnishes the impetus for an examination of the issues raised by framing. His discussion of questions of distributive justice, along with recent experimental results, lead us to propose some selection criteria for which preferences ought to be taken into account in justifiable social choice mechanisms. In particular, an impartial perspective seems a strong candidate as an appropriate viewpoint for eliciting preferences. Finally, we consider a few implications of the discussion for evaluating political institutions.

1/ This paper was written for presentation at the American Political Science Association Annual Meetings, Washington, D.C. Sept. 1997, for a panel on Kenneth Arrow in the series on Nobelists in Economics. The panel was postponed, and the paper was presented at a panel in New York City, November 13, 1997, at the Russell Sage Foundation. Submitted for publication in Taking Economics

Seriously: Conversations with Nobelists about Economics and Political Science. James Alt, Elinor Ostrom, and Margaret Levi (eds.). (Russell Sage Foundation, New York).

2/ We are deeply indebted to Michael Cain for his extensive and most helpful comments of an earlier draft. We are grateful to others including Keith Dougherty, who also gave us helpful comments.

## TABLE OF CONTENTS

ABSTRACT i

SOCIAL CHOICE 4

THE NECESSITY AND CENTRALITY OF ORGANIZATION IN HUMAN LIFE 7

CHARACTERIZING PREFERENCES AND THE EVALUATION OF SOCIAL CHOICES AND INSTITUTIONS  
8

CRITERIA FOR PERFORMANCE BASED ON AGGREGATE WELFARE 10

PARETO (11); BEYOND THE PARETO PRINCIPLE (12)

WHAT ONE GETS DEPENDS ON WHAT ONE'S GOT: ARE ORDINAL PREFERENCES THE PROBLEM ? 14  
FOR VOTING (15)

WHAT HAVE WE GOT TO AGGREGATE? 16

THE CONTENT OF PREFERENCES (17); PREFERENCES OVER THE UNIVERSAL SET (19)

WHAT *MIGHT* WE HAVE AND WHAT CAN IT GET US? 21

PROBLEMS IN IDENTIFYING THE UNIVERSAL SET (21); CONTEXT DEPENDENT PREFERENCES  
(24); MODELS, FRAMING AND CUES: (26)

HOW MIGHT FRAMING AFFECT A SOCIAL CHOICE PERSPECTIVE? 31

INTEGRATING DIVERSE PERSPECTIVES: THE EMPIRICAL CONNECTION (33); A FEW  
IMPLICATIONS FOR POLITICAL DISCOURSE (38)

REFERENCES 41

ENDNOTES 51

## FIGURES AND TABLES

Table 1: Pareto & Majority Rule (61)

Table 2: Outcomes and Pareto (61)

Table 3: Relative Frequency of Outcomes (61)

**Figure 1:** The traditional economics model relating choice to preferences (62)

**Figure 2:** The generation of different preferences in a choice context by cues (63)

## *Kenneth Arrow, Welfare Aggregation and Progress in Political Theory*

by

Norman Frohlich and Joe Oppenheimer

Social Choice and Individual Values (hereafter cited as SCIV), constitutes a fundamental contribution to the social sciences. It is strategically located at the intersection of economics, politics and philosophy. In it, via an abstract and formal model, Arrow identifies basic problems in aggregating individual preferences. He shows that any function, algorithm, or constitution used to aggregate individual preferences to achieve a sensible and democratic group choice runs into immutable difficulties.<sup>1</sup> The problems assume major normative importance when one interprets the exercise, more generally, as attempting to define a method for arriving at group welfare by aggregating the welfares of the individuals.

Although Arrow's work is often interpreted somewhat narrowly as dealing with the construction of democratic institutions to yield voting outcomes, most scholars believe that its scope is much broader. Here, we focus on some implications of his work for the general justification of democratic political institutions, for policy evaluation, and for our understanding of social welfare and justice. These are, we believe, pivotal elements in Arrow's work which parallel the traditional concerns of political philosophy.

Arrow is interested in considering the existence of an empirical mechanism (or at least some possible attributes of such a mechanism) for aggregating the desires (or welfares) of citizens into an outcome with desirable normative properties. The centrality of this to democratic theory should be obvious. The link between democratic theory and individual welfare is via individual choice. If this choice is fixed by preferences in some stable fashion, as presumed in traditional economics, all is well and good. But Arrow introduced the concept of what he called the "universal set:" the set of all possible alternatives, including hypothetical ones, which a person might face. In doing so, he implicitly raised the problems of the nature of preferences and their possible instability. To the degree that individual preferences are, themselves, unstable, or ill-defined, problems for democratic theory abound. With instability, the link between welfare and individual choice is broken; and with that rupture, the justifiability of the link between individual choice and social welfare is raised anew. Issues such as these lie at the heart of Arrow's concerns, over the last 35 years. In addressing these issues our discussion will pay particular attention to a few problems at the heart of classical political philosophy: the fundamental nature of individual values, their relationship to individual choice, and the relationship between those choices and group welfare.

In 1951, when Kenneth Arrow first published SCIV the discipline of Political Science was on a path toward an inductively based social science: a path with deep roots in a humanist-legalist tradition.<sup>2</sup> Although the normative theory of democracy was a continuation of the debate initiated by Plato in his discourses on the nature of the good state, theorists had come to realize that knowledge about institutions needed to be integrated with an understanding of political behavior to explain political choices. The methodology of the field was consistent with those antecedents and was rarely rigorous, in the scientific sense. Explanation was largely the construction of convincing verbal arguments based on relatively unstructured observations.

Social Choice and Individual Values was not of the same genre. The methodology was formal and deductive and, as such, was a close cousin to mathematics. In an unremarked upon irony, the demonstration of a general impossibility theorem was to help establish the possibility of formal arguments about political phenomena. Along with two other works in economics, von Neumann and Morgenstern's Theory of Games and Economic Behavior, and Black's The Theory of Committees and Elections, SCIV demonstrated the applicability of formalized theoretical argument to the understanding of a broad range of political phenomena. In so doing, the three works set the stage for a partial reintegration of political science and economics by showing that one paradigm could perform yeoman service in two disciplines.

The three books addressed different phenomena but shared a number of characteristics. Each was methodologically individualistic: the starting point was the individual. Each posited the same behavioral assumption regarding individual choice: ordinal rationality.<sup>3</sup> And each used rigorous deduction from these and other contextual premises to derive conclusions about what was either possible, desirable, or expectable, with regard to group decisions.

As might be expected, this formal approach was not immediately embraced by political scientists. Some of the reasons for this can be found in a later work of Arrow's: The Limits of Organization. He notes there that:

A ... key characteristic of information costs is that they are in part capital costs; more specifically, they typically represent an irreversible investment. ... (C)odes ... have to be learned in order to receive messages; the technical vocabulary of any science is a case in point. (Arrow, 1974: 39 - 40).

At the time, only a handful of political scientists were ready to make the capital investments necessary to understand the technical vocabulary of the arguments.

Arrow, von Neumann and Morgenstern, and Black had developed theories. Fortunately, theories do not behave like money. According to Gresham's Law bad money drives out good; but good theories drive out bad ones. Green and Shapiro (1994, p. 3), who take quite a different position on which theories are good and which are bad, note that as late as 1957, there were virtually no articles in the American Political Science Review relying on formalized theories posited on rational choice. By 1994 about 45% of all articles used this approach.

What accounted for this conversion? Why did a substantial subset of political scientists ultimately find themselves attracted to this theoretical approach? A brief look at the structure of the theoretical arguments and the substance of the findings in Social Choice and Individual Values can yield insight into why the approach gained converts.

### **SOCIAL CHOICE**

Substantively, Arrow demonstrated that there exists a contradiction among a minimal set of plausibly desirable properties (desiderata) of any mechanism (or rule, or constitution) for aggregating individual preferences into a social choice. That is, he showed that it is impossible to find a mechanism

which can insure the aggregation of individual preferences into a social choice which was transitive, unrestricted in domain, positively responsive to the preferences of individuals, independent of irrelevant alternatives, non-dictatorial and decisive. We can see how each of these properties are desirable.<sup>4</sup>

Three of the properties are directly related to what we consider to be the normative properties of democracy. Informally:

1. *Non-dictatorship* - rules out the possibility that one individual could decide all the social choices;<sup>5</sup>
2. *Positive Responsiveness* - requires that as an alternative gains support from additional people relative to a second alternative, the application of the constitution does not lower the first alternative in its ranking in relation to the second.
3. *Universal Domain* - insures that the constitution or decision rule be able to apply to the full range of possible patterns of preferences held by the citizenry as an input.

Three other properties help insure that the collective decisions of the group be systematic and consistent translations of individual preferences into social choices. (i.e. that the similar individual preferences lead to similar collective choices):

4. *Independence of Irrelevant Alternatives* - requires that the choice between two alternatives, say  $\alpha$  and  $\beta$ , not depend upon preferences of the individuals over a third, not present, alternative such as  $\gamma$  (see footnote ?).<sup>6</sup>
5. *Transitivity* - requires that if one outcome,  $\alpha$ , is socially preferred to another,  $\beta$ ; and  $\beta$  is in turn preferred to a third,  $\gamma$ , then  $\alpha$  must be socially preferable to  $\gamma$ .<sup>7</sup>
6. *Decisiveness* - demands that the decision rule be able to map the preferences of the individuals into a particular collective decision without ambiguity.

His general impossibility theorem demonstrated that any reasonable mechanism for reaching a centralized social decision on the preferences of the group of affected individuals could not satisfy these properties. If the rules were democratic, they could not prevent instability or inconsistency in the group's choices.<sup>8</sup> In any constitution, one of or more of these desirable aspects of aggregation had to be sacrificed.

For example, if a decision rule satisfies all conditions save transitivity, one can, under specifiable conditions, expect to find voting cycles.<sup>9</sup> Political scientists have tended to focus on this potential intransitivity of democratic voting procedures as the crux of the "Arrow problem."<sup>10</sup> They have largely framed the social choice literature as a generalization of the Condorcet voters' paradox rather than looking to the broader problem of aggregating individual preferences into social welfare.<sup>11</sup> This has led political scientists to characterize the theory as dealing primarily about *voting*. But voting, as an individual act involving choice, is only one sort of aggregation (even of preferences), and Arrow's agenda was much broader.<sup>12</sup>

Since the existence of voting cycles were the prediction of a theoretical construct, political scientists had three basic alternatives. They could either look for instances of the anticipated instability;<sup>13</sup> they

could identify institutional mechanisms which violated one or more of the desiderata of a democratic decision process and prevented the instability;<sup>14</sup> or they could attack the assumptions of the model.<sup>15</sup> Theorists, unhappy with the dark side of democracy revealed by the general impossibility theorem, sought to find alternative assumptions and definitions which would allow for less pessimistic conclusions. On the other hand, Arrow's theoretical result cast a clear light on the function of certain institutional arrangements and this allowed empiricists, to sharpened the focus of their search. The field of social choice theory was born.<sup>16</sup>

Because of its severe implications regarding the impossibility of constructing democratic political systems with acceptable properties, the field of social choice has been a thorn in the side of political scientists. For political analysts, the literature defined and confined the engineering problems of voting and constitutional systems in manners which were neither expected, nor welcomed. While a few political scientists and others debated the relevance of Arrow's dissertation, (Dahl, 1956; Coleman 1966) most continued on their way, talking about the issues of democracy as if he had never published. And even though this has changed somewhat over the last 50 years, the implications of the original "impossibility theorem" have yet to be integrated by the scholars in political philosophy: the field which has most at stake. The vast majority of political philosophers, explicitly concerned about the socially 'good' and the socially 'just,' have integrated neither the methods nor the substance of the social choice perspective into their discourse. This is not only because the methodology is foreign and difficult, but also because Arrow's impossibility theorem was initially interpreted as concerned with voting.

#### **THE NECESSITY AND CENTRALITY OF ORGANIZATION IN HUMAN LIFE**

Certainly, Aristotle was correct: we are communal animals. Our lives are very different from those of animals with meager social lives.<sup>17</sup> The circumstances of our survival and our limitations as individuals appear to require collective action to generate social products.<sup>18</sup> But collective action involves problems of social and political organization. In this context, the *social task of politics can be thought of as the activity required to generate collective or centralized decisions for a group of individuals to achieve an optimization of group welfare given the context.*

#### **CHARACTERIZING PREFERENCES AND THE EVALUATION OF SOCIAL CHOICES AND INSTITUTIONS**

This teleological definition of politics gives us a useful starting point. At least since Aristotle's Politics, the justification for government action has been based, in part, on its impact on the welfare of the ruled. Any serious attempt to evaluate government actions from this perspective requires that we be able to compare the welfare of a group of individuals, in the aggregate, in a variety of real and hypothetical situations. From a consequentialist point of view, to say that a decision leading to one state of affairs is better than an alternative decision leading to another state, requires a comparison of the *aggregate* welfare of the individuals in the two outcomes. This is a requirement if one wishes to evaluate either the outcomes of the decisions or the rules and structures which are used to generate such decisions.

The problem with aggregating *anything* is that any objects to be aggregated have disparate qualities. So, for example, in calculating one's networth, one must aggregate the values of different

types of assets. Luckily, there is a common measuring rod for the exercise: market values, calibrated in some monetary unit.<sup>19</sup> As anyone who has packed a family picnic knows, "One can't add apples and oranges," is only partially true: they can be considered as  $n$  pieces of fruit. But how much is the calculated total of a banana and 3 ounces of raisins? Is it 4 portions of fruit? Or (assuming about 80 raisins per ounce) is it about 241 pieces of fruit? Or is it 7 ounces of fruit? And finally, retreating to the market, can we say it is 60 cents worth of fruit? In other words, what we use as the basis for comparing disparate items is determined by our goals, perspectives, and the cognitive tools available to us at the time. Our evaluations are both teleological and based on how we habitually evaluate objects. What are we interested in and accustomed to dealing with: weight, portions, market values or whatever?

Of course, the aggregation problem can get tougher. With fruit, it was sensible to aggregate the disparate items along any of the dimensions mentioned. Weight, value, portions, or numbers of items all can be totaled via a common valuation unit. But this may not always be possible using commonly available units of measure. Take a two child family: with one child in good health, and one sick with a stomach ailment: what is their aggregate state of health? Or even more complicated: what is their aggregate welfare?<sup>20</sup>

At a very direct level, Arrow's social choice argument asked questions about the very existence of such an aggregation: is it a sensible notion? That is, can we find reasonable ways of aggregating individual welfares and if so, exactly what are the conditions which permit the aggregation.<sup>21</sup>

It is only a small intellectual leap from the evaluation of a social decision by its aggregated effect on individuals' welfares to the normative support of democracy. This step requires two conjectures. First one must believe, as did Aristotle, that the affected individuals are, in general, the best judges of their own welfare and, second, that the separate judgments of the individuals can sensibly be aggregated. That is, it requires that some 'vote counting' process (very broadly understood) can be expected to yield a sensible indicator of aggregate group welfare. At this second, very direct level, Arrow's social choice argument demonstrated deep difficulties with the normative properties one could expect from *any process* which aggregated individual choices to reach a common group decision. Since one aspect of the justification for democracy is the existence of a set of sensible aggregating procedures this is a major line in Arrow's work. Indeed, it was this aspect of his work, the problem of finding an acceptable procedure for aggregating the preferences of individuals which was most remarked upon and followed by those who took Arrow's lead.

As indicated above in the examples of adding up fruits, any aggregation requires an agreement on a common unit of measurement. Arrow was initially concerned with aggregating preferences which contained purely ordinal information. Such preferences allow one to answer questions such as "Am I better off now than I was last year at this time?" They also restrict any interpretation of how much better off one is. Ordinal preferences also preclude the possibility of making certain kinds of direct interpersonal welfare comparisons. Arrow considered whether one can characterize what constitutes the best outcome for society in terms of an aggregation of ordinal preferences.

#### **CRITERIA FOR PERFORMANCE BASED ON AGGREGATE WELFARE**



The major program of the economists who deal with the public finance/welfare economics conundrum, as shaped by Arrow, has been to identify and justify institutional structures which generate acceptable outcomes based on individual welfares. In that task, *preferences* have been the *vehicles* to be used in getting from individual values and welfares to social prescriptions.

PARETO: Economists have long agreed that, at the level of the individual, satisfying individual preferences is an acceptable surrogate for achieving individual welfare satisfaction.<sup>22</sup> And in judging such satisfaction when more than one individual is involved, the economists' principle of choice has long been the Pareto principle. If, in a given state of affairs, at least one person can be made better off, while no-one is hurt, one is not at a Pareto optimal point.<sup>23</sup> It follows that the Pareto optimal set is the set of points in which no one individual can be made better off without harming at least one other. The Pareto principle for welfare judgements, then, is that an acceptable choice must lie in the Pareto set.

This makes the Pareto principle one obvious criterion for the evaluation of institutional or organizational performance. Of course the content of the Pareto set depends upon the preferences (welfares) of the individuals. But which, if any, items in the set are *achievable* also depends upon the rules by which the individuals relate. To see this, consider the following example.

Three individuals, Messieurs *i*, *j*, and *k*, live by a stream. They would like to build a bridge across the stream, but the only feasible place to build the bridge is at a spot directly in front of the home of *i*, thereby diminishing both his view and his privacy. As a result, *i* does not support the bridge.

What does the Pareto principle tell us about the status quo? One can't build the bridge without hurting *i*. Hence, an argument that one ought only to advocate or prescribe change which doesn't hurt anyone leaves one committed not to advocate building the bridge.

But other aspects of the example can be fleshed out to change this judgement. *Suppose that i, j, and k are permitted to compensate one another (i.e. make side payments) for any losses one of them may suffer from any choice which affects them jointly.* Then the outcomes in the Pareto set can change. Imagine that the gain from the bridge to the other 2 individuals more than offsets the loss to *i* in the sense that *i* could be given something by *j* and/or *k* which *j* and *k* would gladly exchange for the bridge and which *i* would find better than the 'no-bridge' status quo. Then they could all agree to the bridge and compensation as making them better off. If compensation is permitted (perhaps by letting *i*'s taxes be tailored to reflect his losses) building the bridge along with appropriate (and acceptable) compensation enters into the Pareto set.<sup>24</sup>

Whether compensation is possible and carried out,<sup>25</sup> or not, can determine the size of the Pareto set. Thus, the Pareto set must be calculated with reference to a well-defined set of alternatives. And the definition of that set would vary as either the objective of the argument or as a function of empirical constraints. As the feasible set changes, the subset of it which is Paretian will also change.

BEYOND THE PARETO PRINCIPLE: Of course, there are other criteria for the performance of institutions. Pareto optimality tells us nothing about distribution, legitimacy or fairness. For example, Pareto optimality can clearly be achieved by an efficient dictator. Such a dictator can merely insure that the policies are optimized for her objectives and the society will surely be in the Pareto set (she will be

made worse off if we move from the status quo). Arrow addressed a deeper problem head on, in **SCIV**: the difficulty of *aggregating individual* welfares, judgements, or values, into a meaningful indicator of *social welfare*, by including procedural desiderata which go beyond the Paretian program.

However, if we rule out dictatorships as many (including Arrow) would, and tack on a few other criteria regarding the minimal reasonable properties of any procedure (such as simple notions of consistency), as indicated above, Arrow has shown we can all too quickly be driven beyond both Pareto and the possible.

A quick and easy way to see how this impasse develops is to consider a simple example of the Condorcet Paradox, in which the Pareto optimal outcomes are, themselves, dominated by another non-optimal outcome when pair-wise majority rule (PMR) is used.<sup>26</sup>

That a Condorcet Paradox, can threaten even Pareto is easily shown. Consider a situation (see Table 1) in which 3 individuals  $\{i, j, k\}$  are voting to choose one of 4 outcomes  $\{a, b, c, d\}$ . Now, item  $d$  is not in the Pareto set (note that all 3 voters prefer  $c$  to  $d$ ).

Yet if we use PMR we can get the single *non*-member of the Pareto set,  $d$ , to be the outcome of a *sequence* of votes. To illustrate consider the results of the following pair-wise contests in Table 2:

If the voting starts on the first line, it will end on the third, with  $a$  as the ultimate victor. If, on the other hand, it starts on the second line, and works down the table, we would in fact end on the fourth line with  $d$ , the outcome which is non-optimal. Insisting on more from democracy: i.e. that the outcome of a sequence of votes lead to a result in the Pareto set might seem quite reasonable, but it can not be guaranteed.

### **WHAT ONE GETS DEPENDS ON WHAT ONE'S GOT: ARE ORDINAL PREFERENCES THE PROBLEM?**

The negative findings noted can move one to wonder what is to blame for the dismal conclusions? One obvious problem is that in all of the results discussed above preferences are restricted to ordinal information (perhaps) about the welfare of the individual.

Specifically, Arrow assumes that preferences, which are (in micro-economics) the 'psychological engine of choice,' contain no more than ordinal information: i.e. the individual's rankings of possible alternatives. As such, any aggregation has to be made of such stuff. As he asserts:

I quickly perceived that the ordinalist viewpoint, which I had fully adopted, implied that the only preference information that could be transmitted across individuals was an ordering. (Arrow, 1950, p. 2)

But ordinal information is comparative in only a minimalist sense. The fact that  $j$  prefers  $x$  to  $y$  does not tell us about the welfare *level* implied by either  $x$  or  $y$ . It is similar to my asserting that I was fatter at age 22, than at age 56, and fatter at 56 than at age 32. You don't have much information there. Certainly you don't know how fat I am, much less whether at 56, I am fatter than my lover was when she was 54! And you aren't even helped in that judgement to learn that she is fatter at 54 than she was 12, etc. When using ordinal preferences, trying to get a measure of social welfare from an aggregation

of preferences is analogous to trying to discover, from comparative age/weight statements of the sort offered above, the width of the mattress we would need were my lover and I to sleep in the same bed.

In evaluating voting situations (e.g. "I vote for Goodwin for Mayor because I prefer him, on balance, to the candidates Black and Johnson") one is often restricted to the kind of minimal information that plagues the social choice literature.

FOR VOTING: The results of a vote clearly reflect only ordinal preference information.<sup>27</sup> How does the process of voting as a choice device yield to analysis under the assumptions of preferences based on strictly ordinal information?

The slender assumptions do not prevent outcomes of certain votes from having ethically desirable properties. So, for example, if we restrict Majority Rule to a vote between only two alternatives, one can guarantee a Pareto improvement (May, 1952). Or, if there are more than two alternatives and we restrict the preferences (and hence violate unrestricted domain, see above p. 4), we can achieve some success. For example, if the majority feels passionately that some alternatives are undesirable, these alternatives will be rejected (Downs, 1957). On the other hand, the cycles which are unavoidable may not be unequivocally undesirable. Miller (1983) argues that although the conditions for democratic pluralism may be difficult to justify in social choice terms, the top cycle (Schwartz, 1986) is Paretian and can be shown to be hit upon at least in every other decision.

Other voting regimes may satisfy other desiderata. For example, the voting literature has developed probabilistic voting models for elections. In those models, usually described spatially, the probability of  $i$  voting for an option,  $\alpha$  over another  $\beta$ , is assumed to go up as the distance between  $\alpha$  and  $i$ 's preferred point decreases. From that perspective, there is an "analogy between market competition and political competition" (c.f. Mueller, p. 214). Many of the normative properties of market outcomes and equilibria carry over into the world of political campaigns. This constitutes substantial progress in the evaluation of voting. But probabilistic behavioral models pose problems of their own as leading to indicators of social welfare. Perhaps their value is best appreciated in the narrow sense of establishing that one might be able to reach social decisions.

As was pointed out, however, Arrow's arguments go to questions beyond voting. Across the wider agenda of social choice (see footnote ?) the bottom line is highly negative. And specifically, as they concern the general issues of welfare judgements, the limitations of the information conveyed from *ordinal*, and *interpersonally incomparable* preferences plays a profound role in limiting what one can conclude from patterns of preferences, and choices.

### **WHAT HAVE WE GOT TO AGGREGATE?**

In making decisions about best policies we presumably would like to take into account the aggregate welfare implications of either the set of choices or the set of reports of individual welfare. It is clear that the information content of any aggregate outcome will depend critically upon the information content of what we aggregate. Can we assume more information is available for welfare judgements than is contained in ordinal preferences which don't support interpersonal comparability?

In one respect the utilitarians were right: *were we able to make full interpersonal comparisons of welfare*, we could aggregate the social consequences of all decisions in terms of welfare consequences. However, without interpersonal comparability, we are stuck with judgements no more powerful than Pareto (Soltan, 1996). So, what exactly is the interpersonally comparative status of the stuff we have to aggregate?

THE CONTENT OF PREFERENCES: In discussing the need to expand the content of preferences, Arrow notes that it may be necessary to extend our conception of preference and, do this by considering *the domain of preferences*. As he put it in an early discussion:

The failure of purely individualistic assumptions to lead to a well-defined social welfare function means, in effect, that there must be a divergence between social and private benefits if we are to be able to discuss a social optimum. Part of each individual's value system must be a scheme of socio-ethical norms, the realization of which cannot, by their nature, be achieved through atomistic market behavior. These norms, further, must be sufficiently similar among the members of the society to avoid the difficulties outlined here. (1950, p. 25)

Hence, early on, Arrow saw that any solution to the social evaluation problem which relied on the aggregation of individual preferences required a (very) broad view of the notion of preferences and individual valuations of alternative social states. He wasn't the only one to feel this way. Earlier Pareto himself, and later Bergson, had argued that one needed to include a broad range of individual judgments in any aggregation procedure to get social evaluations (Arrow, 1973b, p. 122). They each recognized that the welfare of one individual might be dependent upon, or a function of, the welfares of others.

So the possibility of a sensible, and justifiable (in Arrow's sense) social evaluation of a social choice *must* be built upon the ethical aspects of individuals' preferences. And the issue must be joined in two ways: "What constitutes the basis of those ethical aspects?" Further we must explore how far such a basis can advance the problem of aggregation. Arrow points out that there are no *a priori* restraints that logic permits us to put on those functions. It is merely the case that an individual's utility may be a non-trivial function of the welfare of others. As such there is, implicitly, some subjective combining of the welfares of others *within* the individual: i.e.  $u_i = f(w_i, w_j)$ . Viewed this way, individual preferences can be thought of as a sort of subjective social welfare function.

But of course, just because there is such a function implicit in the preference structure of the individual does not endow the preferences with enough ethical content to allow it, alone, to be used as a justifiable basis for a social decision. Individuals would certainly be partial to situations which made *them* better off. It is therefore useful to remind ourselves that, virtually from the beginning, philosophers (certainly including the utilitarians, Arrow 1973b, p. 123) urged that ethical content and judgements be based on some form of *impartial reasoning*. As far back as the 1st Century B.C., Publius Syrus of Rome noted that when disputes arise, there is an inevitable problem of bias. His dictum to avoid this problem and get a fair settlement was, "No one should be judge in his own case."<sup>28</sup> But nothing of this sort has been built into the argument. The ethical content, or the appropriate weighting for the other individuals' welfares, is left outside this argument (see Arrow, 1973b, p. 124).<sup>29</sup>

Theories of social justice require more than a subjective preference structure. They require an interpersonal comparison of welfare. Arrow agrees that we must take that step if we want to make judgements regarding social justice. That is, there must be some comparative information which is meaningful *interpersonally*. (Sen, 1973, argues this compellingly, especially in Chapter 1, but also in Sen, 1970, Chapter 9 and 9\*.) Thus, for example, it must, at a minimum, be interpersonally meaningful to say that it is better to be Jack under circumstances  $\alpha$  than Jill under  $\beta$ . Sen argues that this is precisely the sort of judgement that justice must be built on: the sort of empathy which allows us to put ourselves in each others' shoes, and make comparative judgements regarding welfare. But precisely what form these judgements can and should take is less clear.

Rawls' (1951, 1971, 1985) notions of justice, for example, require that a relatively limited interpersonal comparison be possible. So, for example, all that is required is that we can select the 'worst off'. Such "positional" theories permit the development of consistent evaluation criteria for social choices, premised only on the notion that we can meaningfully transmit such positional information when we aggregate choices. (See d'Aspremont and Gevers, 1977; Arrow, 1977b.) This requires that we must expand preferences to cover such comparisons as  $i$  prefers having the attributes of  $i$  under conditions  $\alpha$  to those of  $j$  under condition  $\beta$ . As Arrow puts it: this means substantially more than the individual  $i$  is able to make the judgement that  $i$  is better off under  $\alpha$  than  $i$  believes he would be as  $j$  under  $\beta$  [or  $u_i(\alpha, i) > u_i(\beta, j)$ ]. Rather it means that it is factually the case that anyone would be better off as  $i$  in situation  $\alpha$  than as  $j$  in situation  $\beta$  [or, now without subscripts:  $u(\alpha, i) > u(\beta, j)$ ] (see Arrow, 1977b, p. 152).

PREFERENCES OVER THE UNIVERSAL SET: In attempting to introduce such additional information about preferences in order to yield more determinate welfare judgments, Arrow (1977b) has made some useful, but not widely remarked upon, distinctions which can help us understand what it is "we've got": i.e., the stuff we are aggregating. To begin with, he points out that preferences, as understood in economics, are assumed to be stable. To underline the distinction he wishes to make he distinguishes between what he calls "tastes," which appear to be unstable, and preferences, which are stable. Preferences are an ordering of all the possible alternatives one might face (sometimes referred to as over the universal set), not just over the alternatives available for choice at a given moment (usually referred to as the feasible set).

To get a feel for the difference between these two sets, consider a wine aficionado going to buy a bottle of Merlot at a small corner store. She may have a clear ranking of all sorts of Merlots beyond both those she confronts in the store and those available with her immediate budget.

As it turns out, this distinction, which may appear to be a quibble, is more than definitional. The insistence that preferences be defined over the universal set permits us to import considerable normative material into the discussion.<sup>30</sup> The universal set contains options we don't often consider. And Arrow uses this to drive to some surprising conclusions:

[A]mong the characteristics which determine an individual's satisfaction are some which are not, at least at the moment, alterable. An individual who is ill can meaningfully be said to prefer being well. If in fact there were some medical means of cure, we would test this preference by asking if

he would purchase the services. Clearly the preference would be there whether or not medicine was useful.<sup>31</sup>

We may suppose that everything which determines an individual's satisfaction is included in the list of goods. Thus, not only the wine, but the ability to enjoy and discriminate are included among goods. It is, in fact, true that only some of the goods so defined are transferable among individuals; others are not. But that consideration enters into the definition of the feasible set, not that of the ordering. If we use this complete list, then everyone should have the same utility function for what he gets out of the social state. This does not, of course, mean that individuals agree on the utility of a social state, since what they receive from a given state differs among individuals. (Arrow, 1977b, p. 159).

### **WHAT *MIGHT* WE HAVE AND WHAT CAN IT GET US?**

The notion that everyone has the same preference over the universal set, appropriately defined, would appear to come close to implying that only one representative individual would be required to evaluate the relative positions of all individuals. That is just a short step from saying that the means for an impartial judgement are within the grasp of any individual. Were any individual put behind a "veil of ignorance" in which she was unaware only of which role in the society she played, any choice of alternatives would, by assumption, overcome the difficulties Arrow identified in SCIV. Since all preference structures would be the same and any representative would be a dictator, the choice would be in the Pareto set. And all individuals, *were they to vote from this impartial point of view*, would presumably chose the same alternative.

PROBLEMS IN IDENTIFYING THE UNIVERSAL SET: Of course, such a conclusion assumes that one's utility function is defined over the universal set which is common to all humanity, and hence that neither one's preferences nor the universal set are a function of one's particular experiences. Arrow acknowledges a possible difficulty with his construct:

"... if your satisfaction depends on some inner qualities that I do not possess, then I really have not had the experience which will enable me to judge the satisfaction one would derive from that quality in association with some distribution of goods. Hence, my judgement has a probability element in it and will not agree with your judgment... (Arrow, 1977b, p. 160).

In commenting on John Rawls, Arrow, 1973, alludes to the difficulty of integrating potentially divergent perspectives on the universal set:

To the extent that individuals are really individual, each an autonomous end in himself, to that extent they must be somewhat mysterious and inaccessible to one another. There cannot be any rule that is completely acceptable to all. There must be, or so it now seems to me, the possibility of unadjudicable conflict ... (p. 114)

The difference Arrow draws between tastes, over the feasible set, as opposed to preferences over the universal set, is an important and incisive one. We are not convinced, however, that it is reflective of individuals as they ordinarily make decisions. Perhaps, Arrow's distinction does not go sufficiently

deep (i.e. close to bedrock - see Popper, 1959, p. 111) to cope with recent observational difficulties about individual tastes/preferences. The assumption of a common utility function flies in the face of modern neuro-biology (see, for example, Edelman, 1992) and, upon reflection, also runs counter to common sense.

The assumption of commonality is also problematic in another way. It assumes a strange ontology: the stable existence of the items in the universal set. If the items in the universal set are also a function of culture (was watching the Simpsons on Channel 5 part of the universal set in 1970, in 1870?), then we must define those items in the universal set in terms of individuals' knowledge of the options they believe to be or can imagine to be available.

We conclude that both the domain of preferences, and their functional form are more variable. Individual preferences must contain many unique aspects and a uniform set of preferences over a universal set is most unlikely. Of course, that need not preclude the finding of some decisional consensus and hence the possibility of sensible aggregation of preferences over some domains. But the search for such consensus to gain leverage over the problems of social welfare will need to involve more than a simple polling of ambient preferences over some carefully defined but commonly understood feasible, or universal, set.

Indeed, we can now say that prior to an evaluation of the ethical, or even "welfare" content of a single preference structure, much less, an aggregation of preferences, some ground work must be done. If there can be no consensus regarding the content of the 'universal set' of options, then one must specify the proper *domain* of alternatives over which the preferences operate.<sup>32</sup>

A common domain would identify the set of alternatives to be evaluated, including all relevant attributes of the alternatives. Arrow assumes that the alternatives, once fully specified, would be viewed from a single, unambiguous perspective and yield an unambiguous preference ranking. But more than agreement on a common domain is required. If more than one perspective is possible on the common domain, and if different perspectives yield different preferences, then the choice of an appropriate perspective from which to assess preferences must rest on a defensible normative element. For social welfare purposes, we argue below, an impartial frame of reference is often needed for a preference aggregation procedure to generate compelling results.

CONTEXT DEPENDENT PREFERENCES: Putting it differently, both Arrow, and social choice theorists following him, assume that preferences are fundamental. As such, they are not affected by institutional and other environmental aspects of the choice environment. This permits social choice theorists to accept preferences as adequate representations of individual welfare. It then justifies the evaluation of social choice mechanisms on the basis of their *preference* aggregation properties. But we question the adequacy of a preference structure as unfiltered datum. Hence, we reopen the problem of evaluation of the social choice mechanism by questioning the domain of the social choice function. Here we touch on some of the characteristics of preferences in light of recent findings in psychology. These findings indicate that preferences are both less "solid" and less fundamental than social choice theorists assume. Ironically, by considering the reasons for preference *instability* we hope to arrive at an empirical

mechanism with some improved hope for the sensible aggregating of preferences and with some desirable normative properties.

Among others, Tversky, Kahneman, May, Quattrone and Plott have produced experimental results which show that individual choices, and hence revealed preferences, violate a number of the postulates of the choice model used by economists. Possibly the most famous example is that developed by (Tversky and Kahneman, 1981). There, the subject's response in a choice situation depends on the framing of the decision as either about a loss or a gain. Two verbally different, but logically equivalent, characterizations of a choice situation lead subjects to display contradictory preferences over two alternatives.

Tversky and Kahneman attempt to explain this apparent contradiction by means of a new model of choice which they call "prospect theory". Their argument is that individuals make different choices in the "same" situation when the situation is "framed" as involving a "loss" as opposed to a "gain." Their theoretical representation of such behavior in "gain" versus "loss" situations uses a value function with a discontinuity in the first derivative at the status quo point.

Many articles have since been written arguing that the effects caught in these and other experiments are more general phenomenon, and play a profound role in choice behavior. We take a slightly more general approach than that put forward in Prospect Theory. From our perspective the variability of individual choice is to be thought of as a result of the use of varying models of evaluation. In the case of the Kahneman and Tversky experiment, many people may have one model for dealing with losses, and another for dealing with gains. Interpreted this way Tversky and Kahneman's diagram tries to encapsulate, in prospect theory, multiple views in a single model. As they argue, situations can be framed as a potential loss situation or as a potential gain situation. We would characterize this as indicating that the models that govern our preferences (hence, behavior) from these two perspectives are different; and it is for that reason that the framing affects our choices.

Although, on first blush our characterization may appear to be a distinction without a difference, we think that a broader reading of the results on preference reversal may put this in a different light. Further, we hope to show that our interpretation may have direct implications for some of the questions Arrow raised in discussion of both the universal set and its implications for social aggregation of preferences.

MODELS, FRAMING AND CUES: Recent evidence from biology affords insight into how the brain makes sense of stimuli. Every concrete situation involves a potentially infinite number of aspects. Approaching any situation, because of our limited information processing capabilities, people focus on a small subset of the aspects of the situation which confronts us.<sup>33</sup> We "make sense" of the situation by means of a subset of our neural network which constitutes a cognitive model of the situation. The neural model is literally groups of neurons, each group roughly representing a concept, linked together in a network that is potentiated when appropriate stimuli are received from the environment (or from internal cognitive activity).



So, for example, walking past a schoolyard at recess time and seeing a group of students surround a girl, one might "understand" the situation either as "a game," or as an instance of dangerous "bullying." Which way one understands the situation determines what affect (value) one attaches to the situation because in the brain the neuronal cognitive complex which represents "game" is linked to the limbic system differently than that which represents "bullying." As such they evoke a differential emotional values when stimulated. What triggers the understanding can vary. For example, a sensory input at the time of the observation could help determine our interpretation. Noticing tears, or a smile, could affect our understanding of the situation. But our understanding might be a result of inputs received at a different time. Recent, or distant exposure to media emphasis on problems of bullying in playgrounds, (or our own past playground experiences as children) might favor one interpretation. By sharp contrast media emphasis on the beneficial effects of recess on building peer solidarity might help determine quite a different "understanding" of one's visual inputs.

What is important to note is that in the situation above, the individual has limited information, and the shifting "cues" and frames can be thought of as associated with learning. And it may be that in many choice situations imperfect information, and hence learning, is associated with the instability of preferences over the choices available. But it is important to note that this potential for interpreting a situation in more than one way is not necessarily a function of not knowing what is "really" happening. So even in the playground example, there may be no unambiguous consensus either among the children, or within a given child, as to whether this really is a game or torment.

Judgements, such as these, can be fragile and at times may be a function of partial information. But fragile judgements, unstable preferences, and the like can stem from other causes. May (1954), observing a high frequency of preference reversals, argued that unpracticed judgements can lead to unstable preferences in a situation even with good information.

Larrick and Blount's experimental studies (1997a and 1997b) of ultimatum games provide examples of framing effects which occur without any apparent learning component. In their studies, small changes in wording led to substantial shifts in the patterns of observed behavior.

Larrick and Blount develop an experiment in which structurally equivalent processes, differ only in their verbal description of the choice options, not in the substantive structure of the choices. They utilize Ultimatum Games involving two persons. Characteristically in such games, one subject is given some money (usually \$10) to divide between the two. The second then can accept the division (in which case the money is divided as prescribed by the offer), or reject the offer (in which case no money is given to either person). In their experiments two treatments were applied. In one, the first person makes an offer regarding the split in the \$10 and the second person can either accept or reject it. In the other treatment the first person makes a claim about the maximum amount she would offer, and simultaneously, not knowing how much the first person is offering, the second player makes a claim about the lowest amount that she would accept. There is a common understanding that *if the two claims add up to more than the total, no payment will be made*. This difference in presentation affects both the average amounts offered and the average amounts accepted. Table 3, below, highlights the major results which illustrate the differences (see Larrick and Blount, 1997b). Preferences about

which divisions to offer and accept are clearly sensitive to the way in which the division problem is framed. Framing effects need not always be a function of learning or bad information.

These effects can be related to our discussion about how framing affects preferences. We would say that humans seem to understand (explain) the world by the use of cognitive models, and the models relate to specific aspects of situations. From the experimental results, it seems that framing the situation as one of coincident offers as opposed to sequential bids and acceptances/rejections evokes different models in subjects' minds. These small changes in emphasis, which leave the basic division problem intact, highlight different aspects of the situation, and/or evoke different models, and hence affect behavior.

Kenneth Arrow (1982) has also written about some of this cognitive research. He summarized, "The drawing of inferences depends then on preconceptions, which may be true or false. The cognitive psychologists refer to the 'framing' of questions, the effect of the way they are formulated on the answers. A fundamental element of rationality, so elementary that we hardly notice it, is, in logicians' language, its *extensionality*. The chosen element depends on the opportunity set from which the choice is to be made, independently of how that set is described. ... The cognitive psychologists deny that choice is in fact extensional; the framing of the question affects the answer."(p. 268)

Of course, as with the case of the observation of the children on the playground, there may be some observational learning which goes on. In those cases preferences and behavior may be a function of cues which are used to "better understand" a situation and may reach an equilibrium value when enough information has been gathered. But in some cases, the instability appears to be more basic (as in the case of the Kahneman and Tversky framing of gains and losses, or the Larrick and Blount examples). In those cases the instability may be caused by the cues and framing, and may have nothing to do with learning. The instability may be inherent.

This cognitive interpretation is parallel to current arguments about cognition in the Philosophy of Science. Giere (1990) argues that scientific explanation conforms to this kind of cognitive processing. For him, explanation proceeds by a particular form of analogy. Explanation starts with the identification of aspects of a class of situations which may be germane to understanding the situation. The scientist then constructs an abstract (possibly mathematical) model. A theoretical explanation consists of identifying a class of phenomena, identifying certain aspects of that class of phenomena, and asserting that the class of phenomena resemble the model in that particular set of aspects and to a certain specified degree of accuracy.

A parallel between this notion of explanation and the notion of valuation can be drawn. Even when a domain of alternatives is developed for the purposes of a normative evaluation, the aspects of the phenomena which are abstracted and built into an explanatory model are not determined a priori or definitively by the specification of the alternative set. Rather, the evaluations are at base, dependent on the neuro-physiological maps of the individuals. Thus, competing explanations and framings of a given class of phenomena exist, each based on a subset of the neural maps which can be called up, and hence leading to models referring to different aspects of the phenomena in question. In a parallel way individuals may place different valuations on the same alternatives as a function of differing cues. The

cues lead to the focussing on particular aspects of the alternatives and hence to the identification (or evoking) of a particular model to interpret and indirectly, to evaluate, those alternatives.

Hence our understanding of a given situation is fluid, and the justification of our actions are predicated on our invoked interpretative models. Our chosen actions (or revealed choices) will then be a function of the different aspects of the situation which (quite literally) come to mind (or are emphasized). It follows that, as Tversky and Kahneman note: "When framing influences the experience of consequences, the adoption of a decision frame is an ethically significant act." (Tversky and Kahneman, 1981, p. 458)

### **HOW MIGHT FRAMING AFFECT A SOCIAL CHOICE PERSPECTIVE?**

Let us spell out roughly how this perspective changes the fundamental notion of how values are related to preferences and choices. Below we sketch some implications for preferences over welfare judgments. Suppose an individual,  $i$ , is to make a choice in a specified situation,  $A$ . The situation defines the feasible set available to  $i$ . The traditional economics model would have it that the individual examines the feasible set, identifies the possible outcomes, associated with the feasible set of actions, and evaluates, via  $i$ 's preferences, the alternatives on this basis (see **Figure 1**).

Our modification of this simple scheme adds a layer to the decision process. It interposes a cognitive model which  $i$  uses to make sense of the situation. Conceivably, different models of the same alternatives are evoked by the perception of different aspects of the situation in question. And  $i$  is likely to attach different values to the alternatives in the different models of situations. These values may, or may not, be expressible in a simple coherent fashion as an ordering over the alternatives,<sup>34</sup> but the expression of them, here called preferences over the alternatives to be chosen, and indeed, the outcomes themselves, are derivative not only of these values but by the way in which the situation is revealed (framed) and interpreted.

The cognitive schema in **Figure 2** can be juxtaposed with the simpler preference model of **Figure 1**. In **Figure 2** it is assumed that the situation defines a feasible set and *a variety of evocable cues*. The difference is that each cue set (or frame) draws attention to a different set of aspects of the situation and evokes a different mental model in  $i$ . Each model leads to a different set of preferences, and chosen behavior. The mapping of the feasible set to preferences, and indeed, to outcomes is now determined by the processing of cues by the individual actor. This processing may be thought of as via the evocation of a set of models, conceptions, or images of the situation by the cues. To return to our earlier example, model  $M_1$  could represent the interpretation of the children's behavior as play and model  $M_2$  could represent the same situation as bullying. The different models invoke different preferences and hence choices over the feasible set: say whether to walk on, or to intervene, etc. These variable preferences are potentially rooted in a (not single valued) set over the alternatives given the same situation.

This leads us to conclude that preferences are probably not founded on bedrock. There are likely to be no uniform preferences over a universal set. Preferences are a function of the specific cultural and individual experiences that determine our mental templates.<sup>35</sup> They are also likely to be filtered by the

lenses we use to interpret situations. The structure of the decision environment itself (e.g. a particular market or some other structure) may call forth a number of images<sup>36</sup> clustered around the item, each with its own set of attachments to the limbic system.<sup>37</sup>

It follows that experience (and hence institutional arrangements including those of preference expression and aggregation) are likely to affect the preferences expressed. This would be consistent with the evidence from psychologists regarding framing, and our interpretation of it. It could give us a theoretical basis for understanding choice (and hence voting) as a probabilistic ordinal expression.<sup>38</sup>

That may help to explain individual behavior, but what then, is to be done with the indeterminacy of the mapping between the value structures of individuals and their preferences over concrete alternatives as they impinge on the questions of social choice? How is one to justify policies and institutions? Clearly, the complication we have introduced calls for some further normative qualification of which preferences and which domains are to be considered as acceptable for this judgement.<sup>39</sup>

Thus our approach leads to two difficulties: "Which preferences of a given individual are to be taken into account in a social aggregation?" and "How can divergent preferences be aggregated in an acceptable way?"

INTEGRATING DIVERSE PERSPECTIVES: THE EMPIRICAL CONNECTION: The potential multiplicity (and hence apparent instability) of preferences within any one individual can now be integrated into the evaluative task in Arrow's social choice program.

Arrow noted that one of the big difficulties in the ethical program is that individual preferences, though containing 'socio-ethical primitives' do not stem from an ethical 'stance.' We have noted that to choose from among divergent perspectives which could yield different individual preferences one needs a fixed point on which to stand. Put another way, if one wanted to tap an ethically valid preference structure of an individual, it would have to be invoked by cues which have the property that they are valid for such a purpose. But this is, after all, very much the same requirement that drove Rawls and his many forbearers in ethics to call for the use of impartial reasoning. This is also the wisdom captured in the everyday instructions given to juries to be impartial - to exclude certain types of messages and appeals.

But inducing impartial reasoning to identify the "right" preferences may still not solve the problem of aggregation.<sup>40</sup> Arrow himself has considered the problem at various levels (see for example, p. 55 of his 1952 essay and p. 87 of his 1969 review of an argument by Tullock). He has wondered what level of consensus is likely to turn up via focused impartial cogitation on the social values of concern. In the end, in a discussion of Rawls he is pessimistic that all the difficulties of aggregation will be overcome, although he feels that some progress is likely:

Rawls assumes that individuals are egoistic, their social preferences being derived from the veil of ignorance. But why should there not be views of benevolence (or envy) even in the original position? All that is required is that there they not refer to named individuals. But if these are admitted, then there can be disagreement over the degree of benevolence or malevolence, and the

happy assumption that there are no disagreements in the original position disappears. (1973, p. 106)

If there were a single preference structure for all humanity then impartial reasoning would lead to no disagreements. But given the caveats entered above, that is too much to expect. Let us examine the problem of the appropriate perspective for an individual to take so that she can access ethically acceptable preferences for a social decision on aggregate welfare.

Above we have noted that a conclusion such as that reached by Rawls regarding the best way of distributing primary goods (leaving aside the issue of the specification of those goods) is that interpersonal ordinal comparisons of a particular sort are required. It means that for any two individuals  $i$  and  $j$ , not only does  $i$  prefer having the attributes of  $i$  under conditions  $\alpha$  to those of  $j$  under condition  $\beta$  but also conversely, and the judgment must be shared by all. It means  $u(\alpha, i) > u(\beta, j)$ . But of course,  $\alpha$  and  $\beta$  are situations open to the subjective evaluation in terms of individual models of the sort introduced above. What is the proper or appropriate perspective to take on  $\alpha$  and  $\beta$  to evaluate  $i$ 's and  $j$ 's relative position under those conditions? First, an impartial point of view would appear to be appropriate. The evaluation should not be colored by particularistic concerns. But how can that be achieved? How can one take an impartial point of view?

Elsewhere (Frohlich and Oppenheimer, 1992 and 1997) we have argued that laboratory experiments in which conditions are manipulated to induce aspects of impartial reasoning among groups of subjects are a means of gaining information regarding preferences over principles of distributive justice and that a similar prescription might be effective in other ethically problematic situations.

There are two aspects of such an approach which recommend it as a way of generating insight into individual and group valuations of alternatives. First, as noted above, an impartial point of view - one that must take into account the outcome for all individuals in an evenhanded manner - is one which has normative weight. Second, individuals, discussing and deciding an issue in a group, possess an advantage over a set of the same individuals cogitating alone and reaching individual decisions. As the most rudimentary common sense tells us: each individual has private information based on individual experiences which is not accessible to the others unless that information is shared. If the common domain is to be well specified and understood, that understanding is likely to be fuller the broader the perspective taken on it. Group discussion is likely to broaden the perspective. Ideally it should reflect the perspective of all humanity, but at the least it should use the perspectives of all those affected by the decisions.

Group discussion in a structured environment is a way of bringing to bear wider information and a variety of perspectives so that individuals can expand their understanding of the domain of choice and refine their evaluations of the rankings of different alternatives in that set. If the ethical decision is sufficiently well specified and if the sharing of information under impartiality yields a preferred perspective, then a unanimous decision based on a common understanding of the common domain (as it applies to that situation) may be obtainable.

A number of teams have conducted experiments to determine whether such a core of shared preferences about income distribution exist and can be evoked under conditions approximating impartial reasoning (Frohlich and Oppenheimer, 1992; Lissowski et. al. 1991; Jackson, 1995; Saijo and Turnbull, 1996; and Oleson, 1997). In fact, significant consensus was obtained in Canada, three locations in the United States, Poland, Australia and Japan. This speaks well for Arrow's original conjecture regarding the existence of a universal preference structure (at least over this limited domain). But it gives only qualified support to the notion of accessing any universal preference structure. It may be accessible under only the most precarious of conditions, and, given the way in which our experiences structure the way we perceive and evaluate the world, there may be wide divergence in individual preferences across other issues.<sup>41</sup> Yet the results may indicate the general existence of a sets of *both acceptable* and *rejectable* options under properly specified conditions.

This perspective gives us language to use in a potentially meaningful talk about interpersonal comparisons of welfare. It is possible to imagine, in principle, laboratory experiments designed to determine comparisons among different welfare states of hypothetical (or real) individuals. There is nothing either simple, easy, or magical about conducting such experiments, and they are subject to a host of caveats and objections, but they do furnish a potential way of gaining insight into what exactly we are able to understand about one another and agree upon.<sup>42</sup> Whether there is, or is not, a single ethically justifiable preference over a set of alternatives in a particular domain, is ultimately, an empirical question. With the burgeoning of new experimental techniques in Political Science and Economics (Kagel and Roth, 1995) the time may be ripe for the application of empirical methods to some of the fundamental questions raised by Kenneth Arrow's works.

A FEW IMPLICATIONS FOR POLITICAL DISCOURSE: It now would be useful to identify some implications of these observations for the evaluation of social institutions.

As indicated in the pages above, we argue for the following as premises for the argument: P1) The quality of governance is best measured by something like the aggregate welfare of the individuals subject to the governing institutions.<sup>43</sup> P2) Democratic political institutions use a voting mechanism to aggregate individual choices (each of which can be thought of as a partial articulation of a voter's preference, and hence a welfare judgement) into a social choice. P3) The preference any individual articulates vary as a function of the model used by the individual to understand the situation (or the choices to be made). Different frames may evoke preferences which assign different weights to the welfare of others. In particular some frames may lead to an inappropriate discounting of others' welfare. P4) Morality requires that the welfare of others play a non-trivial role in one's preferences.

This leads us to conclude that: C1) some individual preference structures have greater moral standing than others.<sup>44</sup> In particular, those arrived at from a narrow and ill-informed perspective, have less moral weight. C2) In a given situation, with no change in political structures, a given set of individuals can generate different welfare aggregations depending upon which preferences have been articulated. Hence, C3) some welfare aggregations will be morally 'privileged'.

We can be a bit more explicit about social factors which may affect the way in which preferences are formed and articulated. P5) the structural elements of the polity can profoundly affect the content of

the political discourse in the society. Moreover, P6) the content of the political discourse in a society will affect the model used by an individual to determine her preferences. And further, as noted above, some structures and attributes of the political discourse in the society more readily generate preferences which have greater moral standing. It follows: C4) Such structures and attributes have greater moral standing.<sup>45</sup>

Put simply, in any democracy, the outcome of the governmental process is sure to be a function of the cognitive models used by the individuals in their understanding and evaluation of the political alternatives. The models used by individuals are dependent on the framing of alternatives. The relative quality, if not the ethical standing, of the political outcomes is likely to be a function of the sorts of frames utilized by the individuals in their decision making.<sup>46</sup> These frames are a function of the citizens' experience in the political arena, as well as in other arenas. Thus the welfare of the individuals in a democratic polity is a function of both the aggregation mechanisms of the polity and the structure and attributes of the political discourse in the society. Freer, broader, and more open discourse, is preferable. Of course, in modern democracies governmental decisions are a function of many individual decisions, and no one frame will occur. But it would appear that a probabilistic weighting of these frames (an expectation) should be possible, as a function of the sorts of political perspectives which are emphasized and utilized in the designing and choosing of public policy.

To make these abstract observations more concrete we can consider some simplified examples. The prevailing American practice of interest group politics is usually justified normatively by the presumption that individual preferences, as they appear in the wild and as they are aggregated via a variety of democratic institutions, constitute an appropriate basis for achieving group welfare.<sup>47</sup> Given our arguments above, however, narrow interest group politics (which in its most extreme form might be characterized as identity politics) runs the risk of basing group demands on tastes or preferences formed on far too narrow a perspective. It is altogether reasonable to assume that broadly based public frames for the debate of issues might yield changes in outcomes.

To take another example, the change in the framing of racial issues in the United States during the 60's changed the sorts of alternatives which were chosen. The same might be said of the provision of publicly funded medical care in Canada in the same era. It was framed as a fundamental right, rather than as a commodity. In each case alternatives emerged victorious, in part, as a result of the way various groups framed the situation and evaluated alternative courses of action.<sup>48</sup> And of course, the political structures of the two countries also played a role in bringing about the outcomes (as well as furnishing part of the explanation for why publicly funded medical care was more fully implemented in Canada).

Thus, the circle of inquiry begun in the 1950's comes around, in our discussion, to some of the same issues current then.<sup>49</sup> Now, as then we are led to conclude that both structure and content matter. The quality of democratic regimes is not solely a function of the formal properties of the voting system and of the vote counting procedures. The informal or civil properties of the discussion space can affect, to a considerable degree, the sorts of evaluations and hence choices which citizens make.<sup>50</sup>

Arrow's insights into the problems of aggregation and the implications from cognitive psychology regarding the nature of the preferences which enter into the aggregation, open new paths, on the long trail scholars have been exploring in their attempts to evaluate governments and political systems. In the early days of political analysis, Plato and Aristotle brought the knowledge of their time to bear on these nettlesome problems. The new insights we are gaining into the way people see, interpret, and make decisions about the world need to be integrated into that old quest. To do this, scholars will have to consider how the contributions of social choice, experimental techniques, and individual psychology cast light on the frames individuals use in understanding and deciding their political fates.

#### REFERENCES

- Arrow, Kenneth J. (1950). A Difficulty in the Concept of Social Welfare. Journal of Political Economy, 58: 328-346. [Reprinted with a preface in Collected Papers of Kenneth J. Arrow, Vol. 1, Social Choice and Justice. Belknap Press of Harvard University Press. Cambridge, Mass, 1983. (pp. 1-29)].
- Arrow, Kenneth J. (1952). "Le principe de rationalite dans les decisions collectives. Economie Appliquee, 5: 469-484. Translated and printed in Collected Papers of Kenneth J. Arrow, Vol. 1, Social Choice and Justice. Belknap Press of Harvard University Press. Cambridge, Mass, 1983. (pp. 45-58). Translated title: "The Principle of Rationality in Collective Decisions."
- Arrow, Kenneth J. (1959) "Rational Choice Functions and Orderings." ECONOMICA No. 5, v. 26 n.s.#102, pp. 121-127.
- Arrow, Kenneth J. (1963). Social Choice and Individual Values, 2nd ed. Yale: New Haven.
- Arrow, Kenneth J. (1963b). Uncertainty and the Welfare Economics of Medical Care. American Economic Review, LIII (5:December): 941 - 973.
- Arrow, Kenneth J. (1969). Tullock and an Existence Theorem. Public Choice 6: 105-112. reprinted in Collected Papers of Kenneth J. Arrow, Vol. 1, Social Choice and Justice. Belknap Press of Harvard University Press. Cambridge, Mass, 1983. (pp. 81-87).
- Arrow, Kenneth J. (1973). "Some Ordinalist-Utilitarian Notes on Rawls's Sense of Justice. The Journal of Philosophy. 70. (May 10): 245-263. reprinted in Collected Papers of Kenneth J. Arrow, Vol. 1, Social Choice and Justice. Belknap Press of Harvard University Press. Cambridge, Mass, 1983. (pp. 96-115)
- Arrow, Kenneth J. (1973b). Formal Theories of Social Welfare. in Philip Wiener, ed. v. 4 Dictionary of the History of Ideas. Scribner: 276-284. reprinted in Collected Papers of Kenneth J. Arrow, Vol. 1, Social Choice and Justice. Belknap Press of Harvard University Press. Cambridge, Mass, 1983. (pp. 115-132).
- Arrow, Kenneth J. (1974). The Limits of Organization. New York: W. W. Norton.
- Arrow, Kenneth J. (1976). "Welfare Analysis of Changes in Health Coinsurance Rates." in The Role of Health Insurance in the Health Services Sector (Richard N. Rosett, ed.) New York: National Bureau of Economic Research: 3-23.
- Arrow, Kenneth J. (1977). "Current Developments in the Theory of Social Choice," Social Research, 44: 607-622.



- Arrow, Kenneth J. (1977b). "Extended Sympathy and the Possibility of Social Choice." American Economic Review, 67 (December): 219 - 225.
- Arrow, Kenneth J. (1982). "Risk Perception in Psychology and Economics," Economic Inquiry, 20: 1 - 9. Reprinted in Collected Papers of Kenneth J. Arrow, Vol. 3, Individual Choice Under Certainty and Uncertainty. Belknap Press of Harvard University Press. Cambridge, Mass, 1983. (pp. 261-270).
- Barry, Brian. (1995) Impartial Justice and Individual Discretion, Oxford.
- Bartlett, John (1980) Familiar Quotations 1980), Fifteenth Edition, Little Brown: Boston, Mass.
- Bernholz, P. (1974) "Logrolling, Arrow Paradox and Decision Rules: A Generalization." KYKLOS, 27, 49 - 61.
- Black, Duncan (1958) The Theory Of Committees And Elections, Cambridge: Cambridge University Press.
- Black, Duncan and R. Newing, (1951) Committee Decisions with Complementary Valuations. London: William Hodge.
- Coleman, James (1966), "The Possibility of a Social Welfare Function." American Economic Review, 56: 1105-1122.
- Coughlin, Peter (1984) "Probabilistic Voting Models," in Encyclopedia of the Statistical Sciences, ed. Sam. Kotz, Norman Johnson, and Campbell Read, Vol. 6, NY: Wiley.
- Coughlin, Peter (1988) "Expectations about Voter Choices," Public Choice, v. 44, No. 1 pp. 49-59.
- Coughlin, Peter and Muelvin J. Hinich (1984) "Necessary and Sufficient Conditions for Single-Peakedness in Public Economic Models, J of Public Economics, 25: 323-41.
- Dahl, Robert A. (1954), A Preface to Democratic Theory. Chicago University Press: Chicago, Ill.
- Dahl, Robert A. (1993) "The Behavioral Approach in Political Science: Epitaph for a Monument to a Successful Protest," in James Farr and Raymond Seidelman, eds. Discipline and History. Univ of Mich Press, Ann Arbor: 249 - 267. (originally in the American Political Science Review, 55 (1961), 763-772.
- d'Aspremont, Claude and Louis Gevers (1977). Equity and the Informational Basis of Collective Choice. Review of Economic Studies. v. 44, (June): 199-209.
- de Waal, Frans (1982). Chimpanzee Politics: Power and Sex among Apes. Harper & Row: New York.
- Downs, Anthony (1957) An Economic Theory Of Democracy. New York: Harper and Row.
- Easton, David (1993) "Political Science in the United States: Past and Present." in James Farr and Raymond Seidelman, eds. Discipline and History. Univ of Mich Press, Ann Arbor: 291 - 310. From: International Political Science Review. 6 (1985) 133-52
- Eavey, C. & G. Miller, "Bureaucratic Agenda Control: Imposition or Bargaining?" American Political Science Review, 78, September, 1984: 719-733.
- Edelman, Gerald M. (1992) Bright Air, Brilliant Fire: On the Matter of the Mind. New York: Basic Books.
- Enelow, J. and M. Hinich. (1984) The Spatial Theory of Voting. Cambridge Univ. Press: Cambridge, UK.

- Farr, James and Raymond Seidelman (1993). Discipline and History, Ann Arbor: Michigan Univ. Press.
- Feldman, Allan M. and Kyung-Ho Lee (1987). Existence of Electoral Equilibria with Probabilistic Voting. Public Choice Meetings, Tucson, Arizona. 3/27-30.
- Frohlich, Norman and Joe A. Oppenheimer. (1997) A Role for Structured Observation in Ethics. Social Justice Research. v. 10 (1, March): 1-21.
- Frohlich, Norman and Joe A. Oppenheimer (1992). CHOOSING JUSTICE: An Experimental Approach to Ethical Theory. Berkeley: California University Press.
- Frohlich, N. and J. Oppenheimer, with Pat Bond and Irvin Boschman. (1984) "Beyond Economic Man." Journal of Conflict Resolution. v. 28, no. 1, March, 1984: 3-24.
- Giere, R. N. (1990) Explaining Science. University of Chicago Press, Chicago, Illinois.
- Green, Donald P. and Ian Shapiro (1994). Pathologies of Rational Choice Theory: A Critique of Applications in Political Science. New Haven, Conn: Yale University Press.
- Grether, David M. and Charles R. Plott, (1979) "Economic Theory of Choice and the Preference Reversal Phenomenon," American Economic Review, 69 (September): 623 - 638.
- Gunnell, John G. (1993) "American Political Science, Liberalism, and the Invention of Political Theory," in James Farr and Raymond Seidelman, eds. Discipline and History. Univ of Mich Press, Ann Arbor: 179 - 201.
- Habermas, Jurgen (1990). Moral Consciousness and Communication Action. Cambridge, Mass: MIT [translated by Lenhgard and Nicholson.]
- Hare, R. M. (1963) Freedom and Reason, Oxford: Oxford University Press.
- Hamada, "A Simple Majority Rule on the Distribution of Income," Journal of Economic Theory, 1973: 243-76.
- Hardin, Garrett (1968). The Tragedy of the Commons. Science, v. 162: 1243-1248.
- Hinich, M.J. "Equilibrium in Spatial Voting: The Median Voter Result is an Artifact," Journal of Economic Theory, December 1977, v. 16, pp. 208-19.
- Honneth, Axel and Hans Joas, (1991) Communicative Action. Cambridge, Mass: MIT.
- Hotelling, Harold (1929) "Stability in Competition," Economic Journal. v. 39 (1), 41 - 57.
- Kadane, Joseph (1972), On Division of the Question," Public Choice, 13 (Fall): 47-54.
- Jackson, Michael and Peter Hill, (1995) "A Fair Share," Journal of Theoretical Politics 7 (2, April) 69-179.
- Landa, Janet. 1986. "The Political Economy of swarming in honeybees: Voting with the wings, decision-making costs, and the unanimity rule. Public Choice. vol. 51, no. 1, pp. 25-38.
- Larrick, Richard P. and Sally Blount, (1997a). "The Claiming Effect: Why Players are More Generous in Social Dilemmas Than in Ultimatum Games." Journal of Personality and Social Psychology, 72: 810 - 825.
- Larrick, Richard P. and Sally Blount, (1997b). The Claiming Effect: Procedural Framing in Ultimatum Bargaining. Paper presented at the Economic Science Association, Tucson, Arizona. September, 1997.

- Ledyard, John O. (1995). "Public Goods: A Survey of Experimental Research," in The Handbook of Experimental Economics, John H. Kagel and Alvin E. Roth, eds. Princeton University Press: Princeton. pp. 111 - 194.
- Lissowski, Grzegorz, Tadeusz Tyszka and Włodzimierz Okrasa. Principles of Distributive Justice: Experiments in Poland and America. Journal of Conflict Resolution v. 35, No. 1, March, 1991: 98 - 119.
- May, K.O. (1952), "A Set of Independent, Necessary, and Sufficient Conditions for Simple Majority Decision. Econometrica, vol.20 (October, 1952): 680-4.
- May, Kenneth O. (1954) "Intransitivity, Utility, and the Aggregation of Preference Patterns." Econometrica, vol. 22 (No. 1, Jan): 1-13.
- McKelvey, R. D. (1979) "General Conditions for Global Intransitivities in Formal Voting Models." Econometrica, v. 47: pp. 1085 - 1111.
- Miller, Gary, and Thomas Hammond (1994). Why Politics is More Fundamental Than Economics: Incentive-Compatible Mechanisms are Not Credible. Journal of Theoretical Politics v. 6 (No. 1, Jan.) pp. 5-26.
- Miller, Gary, (1992) Managerial Dilemmas: The Political Economy of Hierarchy Cambridge.
- Miller, Gary J. and Joe A. Oppenheimer (1982),"Universalism in Experimental Committees." American Political Science Review, 76 (2, June): 561 - 574.
- Miller, N. R. "Pluralism and Social Choice," American Political Science Review 77 (1983): 734-747.
- Mueller, Dennis C. (1989). Public Choice II. Cambridge Univ Press, Cambridge, UK.
- Oleson, Paul E. (1997) "An Experimental Examination of Alternative Theories of Distributive Justice and Economic Fairness." Paper presented at Public Choice, San Francisco, March.
- Olson, Mancur, (1965) The Logic of Collective Action. Harvard University Press: Cambridge.
- Oppenheimer, Joe (1975) "Some Political Implications of 'Vote Trading and the Voting Paradox: A Proof of Logical Equivalence:' A Comment." American Political Science Review, v. 69, 3 (Sept.) 963-969.
- Plott, C. (1976) "Axiomatic Social Choice Theory: An Overview and Interpretation," American Journal of Political Science, XX, (no.3, Aug) 511-594.
- Popper, Karl (1959) The Logic of Scientific Discovery. New York: Harper and Row. [First Edition, 1934.]
- Quattrone, George A. and Amos Tversky, (1988) "Contrasting Rational and Psychological Analyses of Political Choice." American Political Science Review. 82 (3, Sept.): 719-736.
- Rawls, John (1951) "Outline for a Decision Procedure for Ethics." Philosophical Review 60, 2 (April): 177-97. Reprinted in J. Thomson and G. Dworken (eds.) Ethics. New York: Harper & Row. 1968.
- Rawls, John (1971) A Theory Of Justice, Cambridge: Harvard University Press.
- Rawls, John (1985) "Justice and Fairness: Political not Metaphysical." Philosophy and Public Affairs. Vol 14 (No. 3, Summer): 223-251.
- Riker, William H. (1982) Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice. Prospect Heights, IL.: Waveland Press.

- Romer, Thomas and Howard Rosenthal (1978) "Political Resource Allocation, Controlled Agendas and the Status Quo," Public Choice 33 (4): 27 - 43.
- Ross, Dorothy (1993) "The Development of the Social Sciences," in James Farr and Raymond Seidelman, eds. Discipline and History. Univ of Mich Press, Ann Arbor: 81 - 104.
- Roth, Alvin, E. (1995) "Bargaining Experiments," in Kagel, John H. and Alvin E. Roth, eds. (1995) The Handbook of Experimental Economics. Princeton, Princeton University Press: 253-342.
- Saijo, Tatsuyoshi, Shusuke Takahashi, and Stephen Turnbull (1996). Justice in Income Distribution: An Experimental Approach. Mimeo: presented at 1996 ISA, San Diego, April 18, 1996.
- Schofield, Norman (1978). "Instability of Simple Dynamic Games." Review of Economic Studies. v. 45: pp. 575 - 594.
- Schwartz, T. (1986) The Logic of Collective Choice. Columbia University Press: New York.
- Schwartz, T. (1981) "The Universal Instability Theorem," PUBLIC CHOICE, 37, no. 3, 487 - 502.
- Sen, A.K. (1970) Collective Choice and Social Welfare. North Holland: New York.
- Sen, A. K. (1970b), The Impossibility of a Paretian Liberal." Journal of Political Economy. vol 78 (Jan / Feb), pp. 152-7.
- Sen, A. K. (1973) On Economic Inequality. Norton: N.Y.
- Sen, A. K. (1977) "Social Choice Theory: A Re-Examination." Econometrica v. 45 (January): 53-89.
- Sen, A.K. (1977b) "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory," Philosophy and Public Affairs v. 6 (no. 4, Summer): 317-344. Reprinted in Jane J. Mansbridge, ed. BEYOND SELF INTEREST. Chicago U. Press: Chicago (1990) (p. 25-43).
- Shepsle K. (1979), "Institutional Arrangements and Equilibrium in Multidimensional Voting Models," AJPS, (February, No.1) 23: 27 - 59.
- Shepsle, Kenneth A. and Barry R. Weingast. "Political Preferences for the Pork Barrel: A Generalization," American Journal of Political Science, 25, 1 (Feb) 1981: 96-111.
- Shepsle K. & Barry Weingast (1981), "Structure Induced Equilibrium and Legislative Choice," PUBLIC CHOICE, 37, no. 3, 503-520.
- Soltan, Karol (1986) "Public Policy and Justice." in Ronald L. Cohen, ed. Justice Views from the Social Sciences. Plenum: New York. pp. 235 - 268.
- Tversky, A. and D. Kahneman. (1981) "The Framing of Decisions and the Psychology of Choice." Science. Vol. 211, January 30, pp. 453 - 458.
- United States, Dept. of Health, Education, and Welfare (1969), Toward a Social Report. Washington: U.S. Govt. Print. Off.
- Von Neumann, John, and Oskar Morgenstern, (1944) Theory of Games and Economic Behavior. Wiley, NY.
- Weingast, Barry (1979). "A Rational Choice Perspective on Congressional Norms." American Journal of Political Science, 23 (May): pp. 245-262.
- Wilson, E. O. (1971) The Insect Societies. Cambridge, Mass: Belknap Press of Harvard University.

## ENDNOTES

- 1/ The logic of his argument has been explored and elaborated upon by many, including, of course, Arrow himself. (See for example, Arrow, 1977; Plott, 1976; Sen, 1970 and 1977; Schwartz 1986.) In this essay we focus primarily on some details of the fundamental behavioral assumptions necessary to develop empirical extensions of Arrow's findings.
- 2/ See Farr and Seidelman (1993), especially the chapters editors introductions to Parts 1 and 2 by the editors and the essays by Ross, Merriam, Ricci, Gunnell, Dahl, and Easton.
- 3/ Of course von Neumann and Morgenstern extended this to develop a weak (i.e. non interpersonally comparable) cardinal utility theory in their elaboration of the behavioral assumptions required to generate certain results. This was a cardinality without much difference for the logic of social choice theory (see Schwartz, 1986). The commitment to the notion that preferences contained nothing more than ordinal information can be seen as one of the major breakthroughs in the analysis, and one of the great stumbling blocks for future work (see below, page 14).
- 4/ Of course, Arrow put forward one set of conditions but quite a few other sets have been proposed, and shown to lead to a variety of similar difficulties. An overview (if somewhat dated), which is particularly clear can be had in Plott, 1976.
- 5/ Broader prohibitions, ruling out dictatorial decision power from sub-sets of the population are shown to be the consequences of similar arguments, with relaxations of transitivity (see condition 5 below).
- 6/ There has been some controversy over the 'proper' specification of this criterion. Most outspoken has been Charles Plott (1976).

7/ The implications of this assumption seem to conform to our notions of what is better and best. But see Sen, 1970, Chap. 1\*, who has an interesting discussion of how limited relaxations of this assumption can be consistent with other interpretations of evaluations of better and best, which however, as Mueller puts it, introduce "a degree of arbitrariness into the process." Thus, although such relaxations spread "dictatorial power across a wider group," (Mueller, p. 389) none of the relaxations get us substantially out of the problem.

8/ Again, many texts show versions of his proof: for an example, see Mueller, 1989, Chapter 20.

9/ This means that in using a democratic decision process, the defeat of some policy A by policy B, and B's defeat by a policy C, might be followed by C's defeat by policy A, leaving no policy preferred to all others and calling into question the justification (on the basis of the normative qualities of its consequences), if not the utility, of the decision process.

10/ of examples of literature in this vein could be cited. The best would have to include such path-breaking works as Schofield (1978), McKelvey (1979), Miller (1983). A few others have noted the problem as of larger concern: See Miller (1992) for an example: he raises issues as to whether there is a metric for strategies for efficiency in a firm when there are cycles. Also, one might look at Riker (1982) for a broader examination of the implications of the argument.

11/ The Condorcet (or voters') paradox shows that using majority rule, transitive individual preferences can aggregate into intransitive collective preferences. Hence, there is no consistency to be guaranteed from a majoritarian system: the same profile of individual preferences can lead to various collective choices. Since, an illustration helps, the reader may wish to examine Tables 2 & 3 on page 61.

12/ Sen (1977) makes a helpful distinction between social choice theory, individual preference and 3

different problems related to the social choice literature: 1. Voting and Democratic Institutions; 2. Social Welfare Judgments, and 3. Measures of Social Welfare. The first concerns the problem of choosing a best candidate or policy based on the preferences of a group. The second concerns what kinds of changes are best for the society, given that some members gain and others lose. This kind of "preference" is a multi-dimensional comparative judgment about social states. The informational basis for this judgment is normally wider than for preferences in case (1) and has implications for case (1). The last concerns measures of poverty, national income, inequality, etc. which have normative implications.

13/ One similar approach was to look for general classes of cases which generated instabilities. One such class is vote trading: See Kadane (1972); Bernholz (1974), Oppenheimer (1975), and the generalization by Schwartz (1981).

14/ Much of the work of Romer and Rosenthal, 1978; Shepsle & Weingast, 1981a & b; Shepsle 1979, Weingast 1979, and others [but see Eavey and Miller, 1984; and Miller and Oppenheimer, 1982 for some disconfirming evidence on their hypotheses] has been oriented toward the discovery of how attributes of the institutional structure induces equilibria when otherwise one might predict cycles.

15/ Here the most successful work has been done by Coughlin et. al. (1984, 1988) but also see Feldman et. al. and Hinich, and the useful discussion in Mueller.

16/ Arrow's methodology and assumptions lured scholars to explain other political phenomena. Anthony Downs (1957) (a doctoral student of Arrow's) extended Black's theory of spatial decision making and Hotelling's (1929) argument about economic competition in a spatially differentiated market into a theory of political party competition - opening yet another field: political competition as spatial

competition (see Enelow and Hinich for a particularly lucid development of these arguments).

17/ We are not alone; there are many other social animals. For example, among bees, the relationship of the individual bee to its hive requires occasional relocation decisions to be made for the population of a hive. (See Landa, 1986; Wilson, 1971; also see de Waal, 1982, for a look at social choices among apes.) Such decisions would appear to be strongly related to the welfare and even survival of the communities which make them.

18/ At a minimum, the family structure implies some collective action. Arrow identifies a number of aspects of our nature which generates the need for centralized or collective decisions. These include a) the needs for public goods, and b) the economies of scale in information (see Arrow, 1974).

19/ But even in that case, different assets carry different levels of risk, and the common measuring rod of current market value represents some aggregation of those risk preferences (not to mention any ambiguity or imperfections in information regarding the "true" value of some assets).

20/ Part of the problem here can be seen to stem from the lack of comparability in health statements. Although one can crudely measure social aggregates of health by such things as average mortality rates, life expectancy, and the like, there is little available for sensible small group or nuclear family comparisons (see US Dept. of Health, Education and Welfare, 1969).

21/ Such aggregation reflects only the ordinal information packed in the statements regarding the welfare of the individual (i.e. that individual is better, or worse, off in one situation as compared to another). Again, there is no lack of indicators of social welfare for larger groups.

22/ Indeed, the link between this aspect of "welfare economics" and the normative theories of the utilitarians is too obvious to belabor.



23/ As Soltan (1986) pointed out, Pareto optimality is what is left of utilitarianism when no interpersonal welfare (i.e. utility) comparisons are allowed. There is an obvious relation between Pareto optimality and the voting rule unanimity: if you can improve some, without hurting anyone, and if people vote to improve themselves, then - theoretically at least - unanimity will lead to change until one is at a Pareto optimal point.

24/ Actually, rationality (and self interest) make even simple situations such as the one sketched much harder to work out in reality than would appear to be the case if there is imperfect information. *i*, after all has an incentive to exaggerate the harm the bridge does him in order to demand higher compensation. Indeed there is an entire area of scholarship devoted to this problem (see the discussion of demand revelation mechanisms in Chapter 8, of Mueller, 1989, and Miller and Hammond, 1994).

25/ The relationship between compensation schemes and social choice is quite direct. Arrow, himself, sketched that relationship in Chapter 4 of his SCIV. There he considered whether there was a system of compensation which could be used as a social choice criteria. On the way he discusses Kaldor - Kicks, Skitovsky, and other ideas for compensation which have been proposed.

26/ By PMR here, we include the property that any item which is defeated may not be reintroduced.

27/ One can argue that, inasmuch as voting is costly, turnout rates may reflect intensity of preference. But that argument is tempered by the further observation that there are reasons for voting other than preference for one candidate over another. (Downs, 1957)

28/ Maxim 545, as quoted in John Bartlett, 1980, p. 111.

29/ Fleming (1952), showed that if individuals each are developing preferences as if they were ethical judges, and if they are "capable of making social welfare judgements for part of the society

independently of the remainder" the individual's judgement will be consistent with a utilitarian ethical template (see Arrow, 1973b, p. 124 and also Hamada (1973) and Sen (1973).

30/ The possibility that the preferences over the universal set could help determine some choices or evaluations in the feasible set caused Arrow to wonder about one of the conditions he imposed in his original argument: independence of irrelevant alternatives. We take a similar perspective below, when we deal with the concept of latent, non-revealed preferences. For when we consider framing as determining a preference over a pair, we are asserting that the preferences are not independent of an (irrelevant) framing factor. As Arrow said later, of the universal set: "In many situations we do have information on preferences for nonfeasible alternatives. It can certainly be argued that when available this information should be used in social choice." (Arrow, 1967, p. 76).

31/ Indeed, recent findings in psychoneuroimmunology seem to hint that entertaining the notion of a cure and desiring it may, in some limited situations which are not yet well understood, affect the immune system and improve one's health.

32/ As ought to be clear from the discussion below, we believe that the definition of the domain may be inseparable from the questions regarding cues and framing. But it isn't clear to us, yet, as to how this aspect of the problem should be handled.

33/ Indeed, a moment's reflection leads one to appreciate that even what constitutes "a situation" requires the selective abstraction of aspects of what confronts us.

34/ Edelman, 1992, explicitly identifies some of the neural structure which must underlie any manifest cognitive (or evaluative) judgment. The ambiguities of multiple maps, and the likely probabilistic

selection of those maps in any decision context is sufficient to generate the deeper structure which we are alluding to here.

35/ This does not mean that there cannot be domains over which shared human experience in different cultures and settings may yield similar values and perspectives. We note below the possibility of agreement about some matters of distributive justice. But we would expect considerable divergence on many matters even within a given culture.

36/ Images, of course, is used metaphorically. The "mental images" may not be visual.

37/ To push the matter even further, it may be a physiological fact that one cannot discover what is the meaning of an individual's preference between  $x$  and  $y$  without furnishing cues. And if one furnishes cues, one may be affecting  $i$ 's preferences over the pairs, and indeed, conceivably over other alternatives,  $s$  and  $t$ , in a subsequent decision. This would be a general limitation related to the irrelevance of other alternatives but also reflecting an uncertainty principle similar to the Heisenberg uncertainty principle.

38/ In light of this, we can reconsider the example above, on page 26, of the children at the schoolyard. There is no reason that the individual passerby needed to see tears to generate one frame or the other. The memories associated with such scenes will generate different affects depending upon what is being tapped either by the scene, or by other thoughts and happenings of the moment. Hence the individual assessor is not likely to have consistent evaluations of the scene were it to repeat itself (and yet catch her attention).

39/ For an example regarding domains, see proposals by Hare, 1963, who ties the types of alternatives to the conditions of impartiality.

40/ A substantial effort was undertaken by the authors to induce impartial reasoning in tests of Rawlsian conjectures (Frohlich and Oppenheimer, 1992). They have been extended to other tasks and generalized (see Frohlich and Oppenheimer, 1997). See also the conceptually related experimental literature (Frohlich and Oppenheimer, 1984 and Roth 1995) which has developed over the measure of self-interest's role in individual preference structures to begin to see how context generates substantive differences in preferences.

41/ Of course, it is possible that although the decisions are consensual, aspect of the frame of the situation, *other than impartial reasoning*, is driving the consensus. But we have tried in our experiments to test for the typical framing effects by running the experiments both as loss, and as gain, experiments, and so on. None of these tests for framing seemed to matter (1992, p. 43 et. seq. and pp. 84-87) It must be noted that the frame, or cues, supplied in the experiments helped determine the result. For example, when Oleson introduced the possibility of unemployment, the distribution of the principles chosen by the groups shifted. Other frames didn't seem to affect the distributions.

42/ Exactly what might constitute agreement, what sampling of humanity might be necessary to reach tenable conclusions, and what conditions would be appropriate are all contentious issues (Frohlich and Oppenheimer 1992, 1997).

43/ This means that the government does not get credit for what it has not accomplished. Of course, by omission, we do not here intend to make light of the rights of and liberties of individuals (which must place constraints on centralized decisions) (Sen, 1970b). Similarly, the distributive pattern of the welfares is an important aspect of the aggregate evaluation. It is assumed by us that much of these concerns will be taken care of by arguing that only some sorts of preferences are morally privileged

(see below). Other procedural aspects of governance, such as fairness, openness, and equality of individuals can not considered here, but are closely parallel to the concerns of democratic procedures. So, for example, May (1952) has talked about anonymity, the property that preferences of I and j will be handled in the same fashion by the institutions, as a desirable property of voting systems.

44/ For example, preference structures induced by impartial reasoning may have a claim to being morally privileged.

45/ In particular, political structures which induce more impartial perspectives can make some claim to being morally privileged.

46/ Sen (1977b) went further. He argued that individuals have the ability to choose their preference structure. For him, they had a responsibility to get beyond an ordinalist metric of welfare, and move on to an interpersonal comparative metric so as to make judgements regarding social welfare. In the argument, he also critiques the notion that we should judge outcomes strictly as a function of the welfare of individuals. He conjectures such a template to be incomplete and that moral judgements need to be introduced in the aggregation process. It is in this spirit that we are proposing that the quality of the socio - political space be evaluated on the basis of the sorts of preference structures which are encouraged to be adopted by the populace in their roles as public citizens.

47/ We leave aside, for the purposes of this discussion the other deep questions of the difficulties of mobilizing large groups with diffuse interests against small groups with concentrated interests (Olson 1967, Downs, 1957).

48/ Indeed, one of the functions of political ideologies is to frame a whole set of situations within a grand interpretive schema so that alternatives disfavored by the proponents of the ideology are

downgraded in the minds of those who buy into the ideology's world view.

49/ After all, it was only a few decades ago that those who sought to understand democracy began to consider the role of informal institutions, see footnote ? and surrounding text.

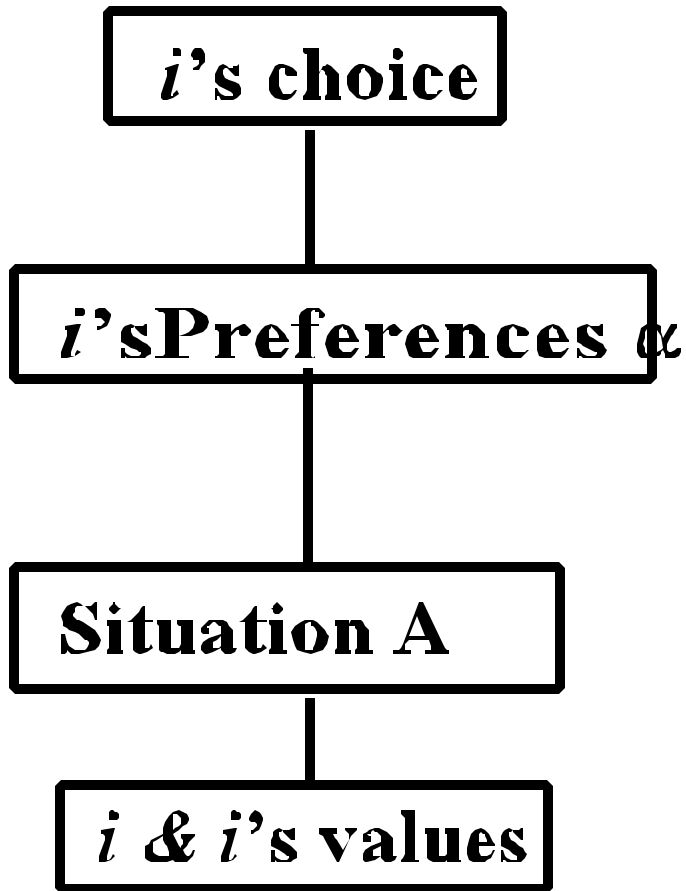
50/ Note how this points toward the perspective of Jurgen Habermas' "ideal speech situation" which may be viewed as an attempt to specify hypothetical conditions conducive to the discovery of important ethical insights. (See Habermas, 1990; Barry, 1995; and Honneth and Joas, 1991). Careful attention to conditions such as these may furnish clues as to how one might structure environments of controlled observation to provide data useful in evaluating ethically "privileged" preferences and hence claims regarding social welfare. If there is validity in the identification of these theoretically identified conditions then by attempting to replicate key components of their constructs we should be able to generate observations that illuminate philosophical issues which underlie the evaluation of democratic systems.

## Tables and Graphs

<b>Table 1: Pareto &amp; Majority Rule</b>			
<b>Rankings</b>	<b>Preferences of the Voters</b>		
	<i>i</i>	<i>j</i>	<i>k</i>
<b>1</b>	<i>a</i>	<i>c</i>	<i>b</i>
<b>2</b>	<i>b</i>	<i>d</i>	<i>c</i>
<b>3</b>	<i>c</i>	<i>a</i>	<i>d</i>
<b>4</b>	<i>d</i>	<i>b</i>	<i>a</i>

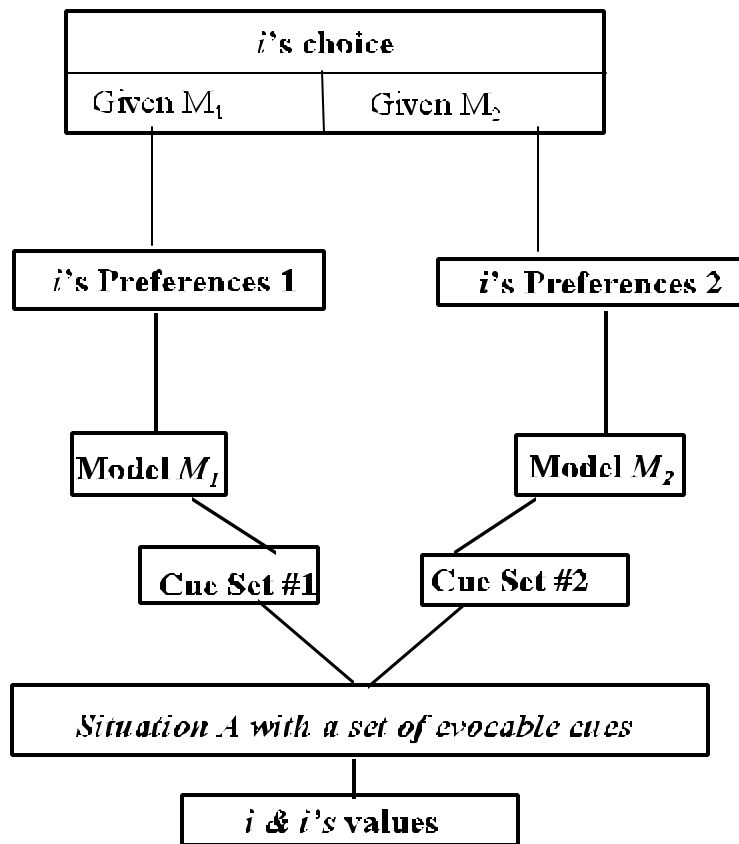
<b>Table 2: Outcomes and Pareto</b>			
<b>Contest</b>	<b>Victor</b>	<b>Voting for Winning</b>	
		<b>Outcome</b>	<b>Vote Tally</b>
<b>d v. c</b>	<i>c</i>	<i>i, j, k</i>	3, 0
<b>c v. b</b>	<i>b</i>	<i>i, k</i>	2, 1
<b>b v. a</b>	<i>a</i>	<i>i, j</i>	2, 1
<b>a v. d</b>	<i>d</i>	<i>j, k</i>	2, 1

<b>Table 3: Relative Frequency of Outcomes</b>		
	<b>Reject / Accept</b>	<b>Claim / Counter Claim</b>
<b>Person #1 offers 50 - 50 split to #2</b>	35 - 50%	55 - 75%
<b>Person #2 lets #1 take all the money</b>	5 - 10%	30 - 40%



**Figure 1:** The traditional economics model relating choice to preferences





**Figure 2:** The generation of different preferences in a choice context by cues