

**DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES**  
**CALIFORNIA INSTITUTE OF TECHNOLOGY**  
**PASADENA, CALIFORNIA 91125**

IMPLEMENTING REPRESENTATION: A FRAMEWORK AND TWO APPLICATIONS

Bruce E. Cain and W. T. Jones



**SOCIAL SCIENCE WORKING PAPER 441**

October 1982

## ABSTRACT

In this paper we introduce a conceptual framework that makes possible a systematic comparison of the many contrasting proposals that theorists have put forward for implementing representation. The exposition of this framework is preceded by a short section in which we discuss various possible approaches to the study of representation and characterize the one we ourselves adopt as "pragmatic," and it is followed by a longer section in which we illustrate the utility of the framework by comparing the proposals of Rousseau and Bentham.

## IMPLEMENTING REPRESENTATION: A FRAMEWORK AND TWO APPLICATIONS

Bruce E. Cain and W. T. Jones

### INTRODUCTION

In this paper we introduce a conceptual framework that makes possible a systematic comparison of the many contrasting proposals that theorists have put forward for implementing representation. The exposition of this framework is preceded by a short section in which we discuss various possible approaches to the study of representation and characterize the one we ourselves adopt as "pragmatic," and it is followed by a longer section in which we illustrate the utility of the framework by comparing the proposals of Rousseau and Bentham.

#### I.

##### APPROACHES TO THE PROBLEM OF REPRESENTATION

In most political organizations the interests of some (the relatively many) are represented by others (the relatively few). The ends of any political organization cannot be realized unless the many are "well represented," that is, represented in ways that further those ends. Because political philosophers and political scientists have differed about ends, they have naturally differed about what it means to be well represented. Hobbes' conception of "authorized" rights, Locke's argument for accountable representatives, Burke's

ideal of the representative as trustee, Hegel's notion of indirect and functional representation, are classical examples of this, the normative, approach to the problem of representation.

But political philosophers and political scientists also differ about how the electorate and its representatives interact. What are the variables -- structural, social, psychological -- that affect these interactions, and how do changes in the variables explain the fact that the electorate and its representatives behave differently in differing circumstances? The attempt to answer such questions is the empirical approach to the study of representation. The Michigan studies of how voters perceive their electoral choice and Fenno's recent work on Congressional home styles are but two examples of the enormous literature that this approach has generated (Campbell et al., 1960; Fenno, 1978).

Finally, since there is usually a gap between the way in which a given electorate and its representatives actually behave and the ways in which that electorate itself, its representatives, or some political philosophers believe they need to behave if they are to be "well represented," a third approach to the study of representation, the "pragmatic," is possible. Aristotle's discussion of how to assure the preponderance of the middle class in a Greek city state, Machiavelli's analysis of the techniques by which a Renaissance prince could maintain himself in power, Madison's discussions of constitutional engineering, Dahl's theory of polyarchy, are examples of this approach -- they are studies of how "desirable" systems of representation can be achieved in various circumstances. Obviously

this third approach supplements the other two: one can hope to bridge the gap between what is and what one believes ought to be only by coming to understand the forces that affect the behavior of a given electorate and its representatives.

In this paper we consider how proposals put forward by theorists who adopt this third, or pragmatic, approach relate to these theorists' underlying assumptions about human behavior and about the ends of the state. The proposals have been so many and varied that meaningful comparisons are difficult and have been seldom attempted. We believe that the framework described in the next section will facilitate systematic comparison and evaluation by bringing to view the underlying logic — and in some cases the illogic — of the various proposals.

## II.

### A FRAMEWORK FOR STUDYING THE IMPLEMENTATION OF REPRESENTATION

Our framework has three components: (1) a goals matrix and (2) a human nature matrix, on which (3) specific proposals for implementing representation can be located and so compared. The goals matrix (G) consists of a set of contrasting values inherent in the way that representatives might act. A profile of the theorist's representational goals is simply a subset of these alternative values. The human nature matrix (B) consists of a set of beliefs, again organized as mutually exclusive alternatives, about human nature. Each individual theorist's view of human nature can be thought of as specified by a certain subset of beliefs in the human nature matrix.

Finally, (3) each proposal (P) can be thought of as a hypothetical imperative of the form: "If you want to achieve such-and-such a G and believe such-and-such B's then do P." Theorists seldom formulate their hypothetical imperative explicitly: more often than not they simply assert "Do P." The framework helps uncover the reasoning that led the theorists to formulate the particular P that he recommends, and thus demonstrates why different theorist put forward different and often conflicting proposals regarding representation.

#### (1) THE GOALS MATRIX

The first component of this scheme, then, is the classification of ways in which representatives might act, or simply, the goals of representative action. We do not of course believe this list exhausts all the possible goals theorists can have for a representative system. We do believe, however, that most important differences in specific proposals can be accounted for by differences in matrix we have formulated. These goals are defined as follows:

#### G1 Short Versus Long Range Time Perspective

Since policies have short and long range consequences, theorists may differ about what the representative's proper time horizon should be. Some issues, for example, may force a choice between immediate benefits and delayed costs, including externalities which will be only realized in the distant future, perhaps by later generations. The designer of a representative system must therefore decide what role short-range and what role long-range considerations

should play and which institutions will encourage what the designer regards as the proper lawmaking perspective.

#### G2 Parochial Versus Common Interest Perspective

A second parameter of representative design concerns the scope of representation. We can think of a continuum of possibilities ranging from, at one end, representatives who respond to the special concerns of particular regions, persons or interest groups to, at the other end, representatives who attempt to do what is best for the good of the community as a whole, however defined. Different theorists will prefer, and seek to set up, representative systems that operate at different points on this continuum.

#### G3 Cooperation Versus Independence

The third dimension concerns the choice of encouraging or discouraging representatives to link their actions with one another's. This is particularly important in cases where there is little or no overlap in voter preferences. The possibility of making bargains, trading off issues of different saliences or forging compromises can be crucial to the function of government. A theorist's position on this dimension reflects whether he wants individual representatives to cooperate in a systematic and regularized fashion (such as in factions or a strong party system) or prefers that they act and think independently.

#### G4 Resistance to Change Versus Readiness to Change

Since change can either be slow and evolutionary or fast and abrupt, and since each type presents a special challenge, theorists differ about what constitutional provisions will best facilitate representatives' adaptation to changes in the environment — for instance, new electoral interests. Accommodation to slow change involves having to cope with the uncertainty of whether change has really occurred: the uncertainty of existence. Representatives accustomed to representing certain interests, values or political factions may become averse to abandoning them even when there is some evidence of erosion in the prevailing electoral cleavage. When fast changes in preferences occur, brought about by cataclysmic events such as the onset of a world war or a depression, the question is whether such changes are permanent or temporary. In other words, will circumstances eventually return to the status quo? Representatives can be very slow in their accommodation to change, or at the other extreme, they can change positions with every shift in public opinion. The constitutional engineer may have preferences about the proper rate of accommodation.

#### G5 Reliability Versus Unreliability

Lastly, there is the dimension of reliability, which refers to the consistency of a representative's promises and actions: when X says that he will do certain things in office, can he be trusted to implement these policies upon election? Whereas the fourth dimension concerns the consistency of a representative's actions over time, the

fifth looks at the consistency of his words and actions at any one point in time. Although it is easier to evaluate the statements of a reliable legislator, thereby facilitating ex ante evaluation by his constituents, at the same time a perfectly consistent legislator may be less able to cope with unforeseen events, or to build coalitions by skillful bargaining. Once again, different theorists will weigh these considerations differently, aiming at constitutional arrangements that reflect their estimate of where the balance lies.

The goals matrix is represented in Table 1.

[Table 1 about here]

## (2) THE HUMAN NATURE MATRIX

The institutions theorists propose in order to implement their ideas about representation are not only related to their goals, but also to fundamental premises about human nature and behavior. These premises may be derived empirically, or they may be assumed a priori, but in either case, they form the rationale for the theorist's expectation that a given institution will produce a particular outcome.

### B1 Human Nature is Basically Good Versus Human Nature is Basically Bad

This distinction has been so widely discussed in political theory that it needs little elaboration here. The pessimistic position holds that power, self interest and greed are the only enduring human motives. Theorists who take this position -- Hobbes and Madison, for

instance, and in the modern period, Morgenthau -- argue for a realist approach to institutions and against overly high expectations about human behavior. They recommend building institutions that take advantage of man's weaknesses (such as pitting self-interest against self-interest to achieve a stable equilibrium). In contrast, other theorists hold that man is basically good and that power, self-interest and greed are no more intrinsic or natural than the opposite qualities. If man is weak or bad, it is only because society has corrupted him. In his designs the lawgiver can and should rely on altruism and/or a concern for the public good.

### B2 Human Nature is Malleable Versus Human Nature is Intractable

Closely related to the first pair of alternatives is the second -- that human nature is either malleable or intractable. The assumption of intractability is most frequently associated with the pessimistic view of human nature: not only is human nature flawed but nothing can be done about it. Attempts to change individuals by legal, moral or educational means are disparaged. The intractability of man's evil state is even perceived by some theorists as an advantage, in that, since individuals always act to protect their self-interests, institutions founded upon these interests have a solid base. Similarly, the assumption of malleability is usually associated with the optimistic view of human nature. If individuals are not born with the right social instincts, then these can be instilled by education or socialization. In fact, from this point of view being "well represented" may be thought to require that representatives and

the represented alike be raised from their self- interested perspectives to a higher public consciousness.

B3 Individuals Can Understand Their "Real Interests" Versus Individuals Understand Only Their "Apparent" Interests

Theorists also differ as to whether or not individuals are capable of understanding their real interests. To assume that individuals can know their true interests is to assume clairvoyant rationality on the part of one or some of the political actors (e.g. the public, their representatives, or at the very least, the theorist himself). This is not simply a rationality of maximization, but one of preference formation as well (i.e. it is assumed individuals have the "right," or rational, preferences). The theorist need not believe that every actor in the political system is perfectly rational in this sense, but only that some actor is and that he will steer the rest of the population in the direction of their real interests. The contrasting assumption is that individuals cannot know their real preferences, or, to put it another way, that they are capable at best only of the rationality of maximization. For theorists of this persuasion the long term consequences of complex actions cannot be known, and they readily admit the relativity of interests over time.

B4 Individuals are Risk Averse Versus Individuals are Risk Acceptant

Theorists frequently differ over whether individuals inherently prefer stability to change, predictability to unpredictability, certainty to uncertainty. These alternatives can

all be grouped under the rubric of differential attitudes towards risk. Theorists who believe that the public is inherently cautious about change, especially those whose own natural preference is the stability, are suspicious of reformers. Burke, for instance, who admired what he thought was the Englishman's pragmatic conservatism, believed that this explained why England was spared the disastrous experience of the French revolution. Others, such as Jefferson or Mao Tse Tung, who also believe that man is innately conservative but who do not share Burke's preference for stability, design measures that facilitate adaptive change, emphasizing for instance the need for bringing about a constant string of small revolutions. Not all, of course, believe that man is innately risk averse: some like Hobbes see man as a restless striving creature with little respect for the status quo. In their view, without the strong hand of the sovereign, society would be in a constant state for chaotic disequilibrium.

B5 There are Transcendental Moral Laws Versus There are Only Secular Laws and Incentives

The implementation of representation will also be affected by the theorist's assumptions about what kinds of laws are relevant to political behavior. If he believes there are transcendental laws, i.e., laws that have a supernatural source — or even if he disbelieves in such laws but holds men in general believe in such laws — he will take them into account in framing his proposals for representation. First, the theorist is likely to hold that the scope and substance of laws of the state are constrained by the condition

that legislation must not violate transcendental principles. Second, he is likely to think that the behavior of representatives is prescribed in ways not affected by institutional incentives.

The human nature matrix is represented in Table 2.

[Table 2 about here]

#### RELATING INSTITUTIONAL PROPOSALS TO GOALS AND BELIEFS ABOUT HUMAN NATURE

How does a theorist's position on the human nature and goals matrices affect his specific proposals for implementing representation? Before we turn to a detailed discussion of Rousseau and Bentham, a simplified example may be helpful. Let us suppose, then, a fictive theorist who is more fully conscious of his positions on these matrices than most theorists probably are. Let us stipulate, first, that one of his goals is to increase the time perspective of representatives in order to provide for longer-range policy planning (G1a) and, second, that he holds the following beliefs about human nature:

- B1b Individuals think in terms of their self interest
- B2b We cannot change this orientation
- B3b Individuals do not know their real interests
- B4a Individuals are risk averse
- B5a No transcendental laws induce a natural concern for future generations

Now suppose that the establishment of a special policy school is being considered — one like the Ecole Normal Superieure, which would be devoted to teaching the tools of long-range planning. How would our putative theorist react? Obviously he should conclude that it is impractical — it would be inconsistent with B1b, B2b and B3b, since they imply that if officials are placed in situations where their immediate interests are short term, no amount of indoctrination will prevent their orientation from becoming short term as well. Thus our theorist would predict that elected officials who must of course constantly weigh the short term electoral advantage of one policy versus the long range good of another would always follow their short term electoral interest, regardless of what they had been taught, and in a sense had "learned," in the proposed policy school.

Only if the theorist held different beliefs about human nature would the policy school proposal appear to him to be practical. For instance, if the theorist believed that individuals can be taught to know their real interests and those of society (B3a), because human nature is malleable (B2a) then he might well conclude that the policy school could succeed in training people to act in the desired way. It would even be possible for him to retain his pessimistic assumptions about human nature (B1b, B2b, B3b) and still justify a policy school if he or she believed that individuals are risk acceptant (B4b) rather than risk averse (B4a). By orienting a policy school curriculum towards showing policy makers how to take risks in the expectation of higher future gain, representatives could be induced to concern themselves about long range policy effects if high, but risky, future



gain could be expected from doing so.

But if a change from B4a to B4b would be enough to shift our fictive theorist from opposition to support of the school, he would certainly point out that it is not enough merely to teach risk accepting people the advantages of risk taking. The institutions into which the school's graduates pass must provide opportunities for such risk-taking. Clearly, therefore, he would conclude that the proposed school by itself would not be adequate to achieve G1a.

Thus a particular institution, say, the policy school, can be justified from more than one set of assumptions about human nature. And if theorists differ about the value of a certain institution, this disagreement may be caused by a difference over goals or over beliefs about human nature. For instance, if one theorist advocates the policy school as a valuable proposal and another rejects it, this discrepancy could be caused by a difference about the importance of short versus long term time horizons, or it could be caused by different beliefs about what motivates individuals.

### III.

#### APPLYING THE FRAMEWORK TO BENTHAM AND ROUSSEAU

We believe that this framework can be used to understand the many ways in which theorists have tried to implement their systems of representation, and we will try to demonstrate its value by using it to analyze the proposals of Jeremy Bentham and Jean Jacques Rousseau.<sup>1</sup> We will first "locate" their goals for representation and their beliefs about human nature on our two matrices, thus making easy a

comparison of important similarities and differences, and we shall then relate these comparisons to their disagreements or agreements about particular institutions.

Locating Bentham's and Rousseau's goals of representation on the Goals Matrix, we see that they diverged on three of these goals G1, G4, G5, and held similar views on one G2. The fifth, G3, as we shall see shortly, is more ambiguous.

As regards G1 -- goals with respect to the representative's time perspective -- Bentham falls closer to the short range and Rousseau to the long range. Since future considerations mattered for Bentham only so far as they could be probabilistically assessed and so far as they affected the present generation, every utility maximizer would necessarily have a short time horizon. He would forego present pleasure to avoid greater future pain, only if the probabilities of alternative consequences could be estimated. An individual's future entered into his calculations only in so far as his present pleasures and pains were affected by the consideration of various future states. Similarly, future generations were to be taken into account by a present generation only in so far as the present generation's pleasures and pains were affected by its thought of future generations. Legislators in the present had no power over or responsibility for future constituents: "The power thus unlimited is that of the legislature for the time being" (CC, p. 160).

Because Rousseau's view of time was very different from Bentham's he differed from Bentham on G1. For Rousseau, time past and time future were "ingredient" (Whitehead's term) in time present.

What was true of a man's personal life -- that he identified in memory with the child he once was and also lived in anticipation with the older person he was going to be -- applied to a society and to the present generation in its relation to past and future generations. Patriotism, for instance, was an emotion binding men to the past of their society and also, and equally, to its future (Poland, pp. 164-65). In sum, the narrowly egoistic and time-bound calculations that Bentham approved as "rational," were condemned by Rousseau as products of "base philosophy, petty self-interest and inept institutions" (Poland, p. 163).

Bentham and Rousseau also disagreed over G4, the importance of representatives being readily adaptive to change. Because Bentham had little concern for the continuity of tradition or convention, he believed that legislators should constantly search for ways to improve existing laws: "the most defective parts will continually tend towards amelioration upon the plan of the most perfect" (Prom. of the Laws, p161).

Rousseau thought that what Bentham viewed as a steady tendency "towards amelioration" was, in societies without a common interest, only the result of the representatives bargaining and jockeying for position as the relative strength of factions changed over time, and so far from being a matter of congratulation, was further evidence of the corruption of most modern societies. The situation was different, of course, where a true common interest exists. Then representatives would be responsive to changes in that common interest, but it was likely to change only slowly, especially if the community were based

on agriculture, rather than on finance and manufacturing, and if most citizens were small-holders.

The third goal on which Bentham and Rousseau clearly disagreed is the importance of consistency between promises and actions (G5). Bentham did not think reliability was important. He viewed representatives as professionals, and the kind of contract that a client has with a professional is not one specifying that the professional carry out his job in a such and such a specific way, but one which requires that he act in a manner that he judges best (Rosenblum, 1978, Chapter 2; CC, p. 145). Hence a representative was responsible only for giving the electorate good reasons how he voted.

Unlike Bentham, Rousseau believed that reliability was important. Rousseau's representative was not, like Bentham's, a professional whose duty was to use his own best judgment; his representative was a citizen whose duty was to reflect the will of his fellow citizens. Accordingly, before the start of every session of the legislature, each representative was to be provided with a set of written instructions, which "have been drawn up with great care," and at the end of each session the representative must give his constituents a full report of his actions during that session. Rousseau viewed this report as having "the utmost importance," for it prevented a representative's actions "from ever being anything but the real expression of the will" of the community he represented (Poland, pp. 193-94).

Whereas Bentham's and Rousseau's goals are rather strongly bipolarized on G1, G4, and G5, the contrasts are less sharp for G3,

cooperation versus independence. Bentham and Rousseau agreed that representatives ought to be responsible to the common interest of the community, not to particular interests but their notions of what a community is differed so markedly as to affect, as we shall see, the specified proposals they made regarding representation. Since Bentham defined the common interest as "the sum of the interests of the several members who compose it," it followed for him that any group, however heterogenous, had a common interest: "in so far as the good of the community taken in the aggregate is the paramount object of the representative's care, no obedience will he pay to any such particular will, to the detriment of what appears to him the universal interest" (CC, p. 160).

For Rousseau, in contrast, the common interest was not merely the aggregated interests of a number of individuals but the interest of genuine community, and not every group, by any means, was a community. Just as time, in his view, was not a series of discrete "nows" set end to end, so a community was not merely a collection of encapsulated individuals. It follows that very large states were not, and could not be, communities; since there was no common interest in such groups, they could not, in the strict sense, be represented -- there was nothing in them to be represented.

Finally, Bentham and Rousseau agreed about one goal, namely the desirability of making representatives independent from factions and parochial interests (G2). But even here they differed in emphasis. Bentham did not worry about the need for mechanisms to achieve coordination; he held that the political market would operate

in the manner of the free economic market. Individuals who sought to promote their happiness without diminishing that of others would find themselves "capable of acting without obstruction." But when an individual's pursuit of his own happiness diminished the happiness of others he would "find obstruction thrown in his way." Bentham was enunciating a version of the Pareto principle: improvements would be made if someone or some group could be made better off without making anyone else worse off. There was no need for special measures to assure cooperation; it could occur on its own:

Thus, then, the principle of self-preference has for its regulator in the breast of each, the consciousness of the existence and power of the same principle in the breasts of all the rest: and thus it is that the whole mechanism is at all times kept in a state of perfect order, and at all times performs to admiration everything it was made for (CC, p. 63).

Like Bentham, Rousseau, was interested in cooperation and in the suppression of parochial interests. And cooperation would occur naturally in communities small enough to have an evident common interest. In larger societies it had to be deliberately cultivated -- here he differed from Bentham -- but not by the kinds of "mechanisms" that Bentham would have favored had he not believed the free market would itself produce cooperation. For Rousseau cooperation would be achieved by cultivating national pride, national patriotism, a national religion, and a national, in contrast to a regional, educational system -- not by external sanctions, but by knowing "how

to direct opinion, and thus to govern the passions of men" (Poland, p. 175).

These then are the similarities and differences between Rousseau's and Bentham's loci on the Goals matrix. As with the Goals matrix, there are more differences than similarities. They clearly disagreed as regards the intrinsic goodness or badness of human nature (B1). Bentham's position here was, in Morgenthau's sense of the term, realistic. Individuals were driven by basic self-interested motives: "Nature has placed mankind under the governance of two sovereign masters, pain and pleasure" (Intro to Prin, p. 1). The legislator had at his command "two instruments -- punishment and reward -- each of which, or both, as in his eyes occasion requires, he employs, in the performance of his work," (CC, p. 35).

Though Bentham acknowledged that there might be individuals who act against their own interests for the sake of some social good -- a kind of false consciousness on their part -- these cases are exceptional. The "self-regarding interest is predominant over all other interests put together" (CC, p. 5), and institutions should be built on the statistical tendency of the many to pursue their self-interest and not on the exceptional behavior of the few: "It is in what has place in the conduct on the part of the thousands, and not in what has place in the conduct of one in every thousand, that all rational and useful political arrangements will be grounded." Once this principle is recognized, he claimed, "all pretence to this species of purity will be regarded as would an assertion of chastity in the mouth of a prostitute at the very moment of solicitation" (CC,

p. 61).

As regards Rousseau, although there are passages in which he sounds every bit as "realistic" as Bentham -- for instance, "The great springs of human conduct come down, on close examination, to two, pleasure and vanity; and . . . in the 1st analysis . . . everything comes down to practically pure vanity" -- that human nature is basically good was a central thesis: Since "everything is good as it comes from the hands of the Maker" (Emile, p. 11), it follows that "man [too] is naturally good" (Emile, p. 106).

To say that man is naturally good -- that he is "born free" (Social Contract, p. 5) -- means that at birth his career is open. In a favorable environment he will develop into a virtuous man and a good citizen: "a young man brought up in happy simplicity is drawn by the first movement of nature to the tender kindly passions"; he has a "compassionate heart"; he is "the most generous and most lovable of men" (Emile, p. 102). Unfortunately man is "everywhere in chains" -- that is, in most European countries the children grow up in such an unfavorable environment that, become men, they "pervert and disfigure everything." (Emile, p. 11).

This brings us to B2, beliefs about the malleability and intractability of human nature. Self-preference and utility maximization were the only true bases for human action in Bentham's scheme; in this sense, human nature was intractable. Attempts to educate men to act altruistically were not only doomed to failure, they were pernicious in the sense that they fostered hypocrisy and pretension. Bentham considered various challenges to the

inevitability of self-preference and dismissed them. Religion, for instance, was a false consciousness (B5a) fostered by "hypocritical knaves" who "with the full consciousness of its absurdity" inculcate a doctrine of "sympathy" and perpetuated by "the miserable dupes" who refused to recognize its absurdity (CC, p. 35).

In contrast, in Rousseau's view human nature is malleable. Indeed, because men internalized the belief and value systems of the institutions in which they lived, human nature was diverse as well as malleable — diverse because the basic drives were mediated and articulated by the internalized value system; malleable because, having been shaped by one set of institutions, they could be reshaped by a different set.

Any theorist who proposed to rely exclusively on "mechanical devices" (Corsica, p. 277) — among which Rousseau would almost certainly have included any Benthamite system of external sanctions — would be defeated by the seeming intractability of human nature and by the power of such passions as patriotism and religious fervor, which often resisted any rational calculus of pleasures and pains. But a theorist who took account of the national character of the people for whom he is designing a representative system and who understood their history would find human nature malleable enough.

Again, there is a significant difference between Rousseau and Bentham, on the issue of whether individuals can know their real interests (B3). Bentham denied that men have real interests which are better for them than their perceived self-interests. Of course he recognized that actions have secondary consequences: the "fecundity"

of pain and pleasure was an important consideration for him inasmuch as laws must be judged by their distant as well as proximate consequences; Bentham was aware of the need to discount for the bias of nearness. But, Bentham was certainly not trying to liberate individuals from their self-interested perspective, nor did he think that anything could be in an individual's interest if it did not make him better off, or at least not worse off.

In contrast to Bentham, Rousseau held that there were "real," in distinction from merely "apparent," interests. Men's real interests were to live simply, in peace and amity with one's neighbors; to "seek satisfaction within oneself rather than in one's fortune"; to "perceive God everywhere in his works" (Creed, pp. 3, 20). That is happiness. That being the case, Rousseau had to ask himself why most people pursue only apparent goods? His answer was, first, that the profit motive has made men "scheming, ardent, avid, ambitious, servile and knavish" — in a word, unhappy — and, second, that there was nothing in the contemporary educational system to counteract these tendencies, for whereas "justice and goodness are not merely abstract terms, moral entities created by the understanding, but real affections of the soul (Emile, p. 105), the educational system unfortunately emphasizes only "book learning," that is, what is verbal and abstract.

Thus Rousseau and Bentham diverged on B1, B2, and B3. On B4 and B5 they agreed. As regards B4, both seem to have believed that individuals are by nature risk averse (B4). Bentham constantly assumed that representatives would seek to make their positions

secure; as he often put it, they would try to minimize their dependence on public opinion. However, Bentham was keenly aware that in order to control public officials it was essential that they be insecure and that their dependence on the public be maximized: "if the possessor of the power is, at all events, to keep his hold of it so long as he lives, or even so long as he remains legally unconvicted of a specific misdeed, the difficulty of dealing with him may be unsurmountable." Accordingly, "those who establish government must begin with establishing insecurity: insecurity, viz. as against those in whose hands the means of security against others are reposed" (CC, p. 58).

As for Rousseau, it follows from the fact that the past accumulates in a people's present, that he thought that most people are risk-averse. It was well, in his view, that they are: nothing could do more harm than for a legislator to ignore these inertias in order to follow some abstract model or to take over wholesale the constitutional practice of some other society with a different history and a different ethos. "I cannot repeat it often enough; think well before you lay hands on your laws, and above all on those that have made you what you are" (Poland, p. 182).

Finally, Rousseau and Bentham agreed that the behavior of representatives would have to be controlled by secular laws (B5). In Bentham's view, appeals to transcendental laws were "but so many ways of intimating that a man is firmly persuaded of the truth of this or that moral proposition, though he either thinks he need not, or finds he can't, tell why" (FG, p. 269). The "natural consequence of such

doctrine is to impel a man, by force of conscience, to rise up in arms against any law whatsoever, that he happens not to like," (FG, p. 287). For Bentham the only true justification of any law whatever was utility.

Though Rousseau sometimes used language that suggested an appeal to transcendental laws, the account he usually gave was empirical and, in the broadest sense, psychological. "The most inviolable law of nature is the law of the strongest. No laws, no constitution can be exempted from this law" (Poland, p. 236). Certainly, it was to the feeling heart rather than to the calculating head that he usually appealed as a source of authority for our actions, but that source is clearly human, not divine.

#### IV.

##### SOME INSTITUTIONAL PROPOSALS BY BENTHAM AND ROUSSEAU

We shall illustrate the general relation holding between loci in the G and B matrices on the one hand, and specific proposals on the other hand by a discussion of six institutional issues that were considered both by Bentham and by Rousseau. Of these six, they clearly disagreed on the value of two (i.e. factitious honors and mechanisms to control the conduct of the legislators), partially agreed on the value of two others (i.e. monarchy and separation of powers) and agreed on the last two (i.e. the importance of public opinion and the frequency of elections). We will consider the following questions: (1) What was the basis for their positions? (2) If they disagreed, was it because of a difference in goals or in

beliefs? (3) If they agreed, were their reasons similar? and (4) Were their reasons consistent with their philosophies generally?

#### 1. INSTITUTIONS ABOUT WHOSE VALUE THEY DISAGREED

(i) Factitious Honors. Bentham strongly opposed the use of factitious honors as a method of rewarding public service, because, unlike "natural dignities," they are conferred by a third party (typically, a monarch), not by those who directly perceive the value of the act being rewarded. It was important to Bentham that rewards not be mediated by third parties: in his words, "affection, esteem and respect, which is the result of judgment which is unperverted by any delusion from source, is preferred to that respect which is the joint offspring of sinister interest, caprice, imposture and chance" (CC, p. 90). Under a system of factitious honors, individuals are likely to be accorded honor by accident of birth, or because they please the monarch, rather than by such utility-maximizing criteria as whether the action fosters aggregate utility.

Again, such awards were a source of great waste of public expense (CC, pp. 80-81), and they unnecessarily aggravate inequality. Bentham knew that some measure of inequality is an inevitable by-product of competition. But individuals who earn their inequality by the merit of their actions actually contribute to the general good: the benefit of the inducement exceeds the cost of reward. But when reward is conferred by a third party, there is no guarantee of benefit to the public — indeed, the opposite is more likely to occur.

The proper method of bestowing honor, therefore, is to keep the public informed with the "utmost degree of clearness, correctness and completeness possible," regarding the actions of representatives and other officials so that good actions will be rewarded with natural dignity -- i.e., dignity that comes directly from public esteem. In short, Bentham opposed factitious honors because they undermined the ability of the representative to foster the common good (G2b) and to act independently of the monarch who bestowed these honors (G3b).

Inasmuch as Rousseau shared these goals, it may be initially puzzling to find that he favored the use of such honors. The explanation is his radically different conception of human nature (especially B1, B2, B3). Since in Rousseau's view the particular pattern of a man's motives is shaped by the value system he has internalized, Bentham's notion of there being an incentive system in which "judgment is unperverted by any delusion from source" was itself a delusion, the product of Bentham's oversimplified view that men are always and only animated by pleasure and pain. All incentive systems, Bentham to the contrary, are mediated by third parties -- that is, by the culture in which the individual lives. The only question is what sort of value system is in fact mediating that individual's incentive system, whether it is socially helpful or socially harmful, and it is always possible, by altering the value system, to reorient a socially harmful incentive system in a socially beneficial direction.

Thus most Europeans lived in a society that valued money above all else, and unfortunately "of all the incentives known to me, money is at once the weakest and most useless for the purpose of driving the

political mechanism toward its goal and the strongest and most reliable for the purpose of deflecting it from its course" (Poland, p. 227). But this did not have to be the case. That "reserve of grand passions" which exists "in all hearts" can be redirected away from greed and toward honor, which is a socially useful motive. "I should like to have all ranks, offices, and honorific awards distinguished by external signs, so that no public figure would ever be allowed to go incognito, but would be followed by the marks of his rank and dignity; this would make people respect him at all times, and to dominate over opulence" (Poland, 229).

(ii) Monitoring Legislators. There is another initially puzzling contrast between Bentham's and Rousseau's attitudes toward monitoring: Bentham stipulated in great detail careful day-to-day monitoring of the behavior of representatives, but he left them free to vote in accordance with their best judgment on all issues that came before them. Rousseau, in contrast, believed that representatives should be bound to vote on all issues according to the detailed instructions they had received from the electorate, but he did not think that the day-to-day conduct of representatives required any monitoring. These differences, we believe, can be explained by reference to their differing positions on the Goals and Beliefs matrices.

As an example, let us take Bentham's provisions for monitoring attendance. Attendance was particularly crucial for him because legislators who did not attend produced half the pernicious effect of bad bills passed in their absence, but were not held responsible

because the cause of their absence could not be known. He therefore proposed that attendance be mandatory. The legislative chamber was to have only one entrance where the legislator would be paid per day by the doorkeeper. The moments of arrival and of departure would be recorded in an "Entrance and Departure Book." Each day's attendance record would be printed in the newspaper the next day, and the monthly totals would also be reported. If a representative was ill, he would have to present a "sickness ticket," attested to by a physician, which would indicate the number of days missed and the nature of the illness. In the case of absence, for whatever reason, the legislator was encouraged to designate a substitute. The legislator and not his constituents should choose the substitute: otherwise he could blame the substitute's votes on his constituents. This was consistent with Bentham's arguments that the legislator was responsible for his own judgment and that he should not be able to duck the public's retrospective evaluation.

There is nothing in Rousseau that corresponds to Bentham's minute recommendations for monitoring the day-to-day behavior of representatives. One reason, as we saw with factitious honors, was his very different view of human nature: so far from sharing Bentham's extreme suspicion of human motives, Rousseau believed individuals could come to know and aim at the common good (B1a). Another reason is their differing goals for representatives -- for instance, their differences with respect to G5. Bentham did not think that it was desirable to make representatives act reliably -- i.e., consistently with their promises (G5b) -- whereas Rousseau did. The



key here is the difference between thinking of the electorate as an aggregate and thinking of it as a community. In the Benthamite system it is essential that there be no statistical biases in the sample of legislators when votes on bills are taken.

If the diverse interests of the aggregate which is the legislature systematically reflect the diverse interests of the aggregate which is the electorate, then the votes of the representatives will, with no more ado, reflect those interests, and the greatest possible good will thereby be achieved. Obviously, monitoring of voting is not required, though monitoring of attendance is.

Since Rousseau believed that a real community of interest is possible and, what is more, that both the electorate itself and its representatives can know what that community of interests is, representatives are "not to express their own private opinions but to declare the will" of the electorate (Poland, p. 194).

Note that though Rousseau's proposal that the citizens "bind the representatives to follow their instructions exactly" (Poland, p. 193) is consistent with his goal that representatives should reflect the common good of small communities (G2b), it is inconsistent with his belief that representatives can be socialized to work for the common good (B2a, B3a), for if small communities can foster civic-mindedness, an identity of interests between the representative and the represented should develop without instruction. Perhaps this anomaly betrays a suppressed but realistic suspicion about the limits civic minded motives.

## 2. INSTITUTIONS ABOUT WHICH BENTHAM AND ROUSSEAU PARTLY AGREED

(i) Monarchy. Whereas Bentham adamantly opposed this institution, Rousseau was willing to make adaptations to local customs and preferences. Bentham's reasons for opposing monarchy in any form or under any conditions were consistent with the basic premises of his theory of representation. The goal of the representative was to maximize the aggregate utility of the whole population (G2b). This requires that all preferences be equally weighted. But in a monarchy, if the preferences of the monarch weigh more heavily, the principle of the greatest happiness to the greatest number is therefore necessarily violated. Far from pursuing measures that increased aggregate welfare, the monarch would inevitably look out for his own interests: "In pursuance of the self-preferences inherent in human nature, the end of his government will be the greatest possible happiness of his individual self" (CC, p. 128).

Further, since, as Bentham held, all individuals base their decisions on a calculation of self-interested costs and benefits (G2b, G3b), and since in a monarchy the citizens are not in a position to inflict costs on the monarch, the monarch in turn had less of an incentive to act benevolently towards them: "The more extensively a man feels himself exposed to ill-treatment at the hands of others, the stronger is the inducement he has to bestow upon them good treatment . . . but the monarch is of all men the one who stands the least extensively exposed to ill-treatment at the hands of others" (CC, p.130).

Rousseau, too, opposed monarchy, both for reasons similar to Bentham's and for reasons of his own that reflect a different position on the Goals Matrix. Because the interests of the monarch almost inevitably conflicted with the common interest of the citizens, the monarch would try to disrupt that general will if it existed or to prevent its formation if it had not yet come into being.

Nevertheless, when he was invited to draft a constitution for Poland he recommended modification of the monarchy and the nobility, rather than abolition.

That Rousseau, unlike Bentham, was willing to compromise with monarchy, follows from their differing positions on the Beliefs matrix. His belief in malleability (B2a), and his reverence for continuity and tradition (G4b) led Rousseau to believe that gradual change is possible, while the principle of risk aversion (B4a) to which he and Bentham both adhered, led him to believe that radical change is dangerous. Changes in the institution of monarchy might not only substantially reduce its danger to the state, but actually turn it into a socially useful institution.

Since Rousseau did "not believe that a state as large as Poland could possibly get along without" a king, the problem for the Poles was to limit the monarch's power to do harm. Election helped, he thought, but did not remove the danger of autocracy. Another way to weaken the power of the monarch was to "prescribe by constitutional law that the crown should never pass from father to son, and that every son of a Polish king should forever be excluded from the throne" (Poland, p. 210). Still another was to let the king continue to

appoint the great officers of state but to "restrict his choice to a small number of nominees presented" to him by the legislature, thus making it impossible for him to "fill these offices with his creatures" (Poland, p. 207). Rousseau also urged the Poles to "minimize, as far as possible, the handling of money by the king" (Poland, p. 208).

But in the long run, and in accordance with Rousseau's constant emphasis on the importance of an internalized belief system (B2a), he held it better to try to change the king's perception of his interests than to rely exclusively on constitutional sanctions to confine those interests. How, then, could the king's perception of his personal interests be changed to coincide with the common interest? Since "only a base soul is insensible of posthumous reputation," the Poles should make use of the king's natural concern for history's judgment of him (Poland, pp. 264-5). A special institution should therefore be established, to come into being immediately after a king's death, to evaluate his reign and to award him monuments and honor if he merited them — or to withhold them. It was important that this judgment be made during the interregnum, lest the new king use his position to influence, and mitigate, the assessment of his predecessor.

(ii) Bicameralism. Here again, whereas Bentham was adamantly opposed to the institution, Rousseau was willing to compromise. Bentham was firmly opposed to bicameralism because he held that if one chamber properly embodied the will of the aggregate, then a second legislative body must be either opposed to that will or redundant. If

a second house is different in composition from the first house, it is usually a permanent aristocratic body such as the House of Lords in England. Such houses act against the best interests of the whole population and in the interests of a privileged few. The argument that the members at the second house would have special intellectual aptitudes did not sway him: cleverness was pernicious when it opposed the will of the people.

On the other hand, if the composition of the second house was like that of the first, unnecessary delays would be caused while the two legislatures passed things back and forth. If the first house was sufficiently numerous, Bentham thought, it would probably make the right decision since the "probability of right judgment will in every instance be in the exact ratio of the number of the majority to the minority" (CC, p. 115). Complicating the legislative process by introducing a second house would only make it easier for individuals intent on mischievous aims to succeed.

Like Bentham, Rousseau believed that ideally there should be but one legislative body, but in many states an upper class already existed, represented by a senate which had large powers because its members were either hereditary nobles or else served for life. This was the case in Poland, and instead of recommending abolition of the Polish senate, Rousseau characteristically proposed a number of changes which he thought would reduce the senate's power and yet enable it to make a positive contribution to the state.

For instance, the right to appoint senators should be transferred from the king to the provincial assemblies. This would

have two advantages: first, the king's power would be reduced; second, Poland would move further in the direction of federalism. In a nation as large and diverse as Poland, Rousseau thought, it was unlikely that a communality of interest could be achieved for the whole nation but only for each of the "tiny states" which, federated together, compose the nation. It seems, then, that he advocated for the whole state something like a balance of powers, with each of the several powers representing one more or less common interest.

### 3. INSTITUTIONS THAT BOTH THEORISTS AGREED ABOUT

(i) Public Opinion. Since Bentham and Rousseau were both aware of the power of public opinion to influence the behavior of representatives, both proposed institutions to focus and direct public opinion. But because of their different positions in the human nature matrix the institutions they recommended differ markedly. Bentham, being concerned about the need to check the honesty and competence of public officials, proposed the establishment of what he called the Public Opinion Tribunal. The duties of this institution were several: it would provide evidence and make judgments on matters of public interest, it would have the power to render or withhold offices and it could make recommendations for improvement. He proposed to admit all who wanted to be members to this tribunal, drawing an analogy with committees in the House of Commons for which there was a rule that all who came to the committee should have voices. At another point he compared the Tribunal's function with that of a jury which oversaw and judged the conduct of functionaries, bringing "the force of the

popular or moral sanction to bear with greatest advantage upon the conduct of public functionaries in the several departments" (CC, p. 42).

It is surprising that Bentham, who was so careful to prevent biases in the sample of preferences in other decisionmaking bodies, such as his proposed Assembly or even the electorate, should not have had a similar concern about membership in his Tribunal. He did indeed express some concern that aristocratic members might have interests that diverged from those of the more democratic members, and warned that when the Tribunal delegated functions to smaller sections, it should carefully limit or exclude the aristocratic members. But the possibility of other sampling biases -- e.g., interest groups -- was ignored. This omission is important, for unlike the legislators, who were instructed to use the calculus of utility to derive the common good, members of the Public Opinion Tribunal were to decide on the basis of their own interest, and the opinion of the group was the sum of these interests. If, as seems likely, Bentham intended the Tribunal as a kind of running Gallup poll, he neglected the various sampling biases one would expect to find, given his assumptions about human nature (B1b, B3b). If individuals are utility maximizers and if there are costs to information, attendance would covary with interest and the decisions of the Tribunal would reflect the intense interests of particular interest groups rather than the aggregate utility of the whole population.

The nearest equivalent in Rousseau's thought to Bentham's Public Opinion Tribunal was the special tribunal which, as we have

seen, he thought should convene on the death of the Polish king to render a verdict on his reign. Since this tribunal was certainly not a running Gallup poll, and since it had a one-time only function, the possibility that sampling biases might affect its verdict was much less importance than in the case of Bentham's Public Opinion Tribunal. What was essential was that the king know ex ante that a tribunal would be convened; it was less important that the tribunal actually reflect public opinion accurately than that the king believe it would make "a just and rigorous judgment of his conduct" (Poland, p. 265).

But in any case, Rousseau's lack of concern about the possibility of sampling biases occurring in the expression of public opinion reflects his view that in a society in which a truly common interest exists, sampling errors simply could not occur, whereas in a society which did not already have a common interest no amount of care in sampling would produce one. Thus whereas Bentham supported the Public Opinion Tribunal as a mechanism for discovering the public good, Rousseau saw his (rough) equivalent of it as part of the socialization process discussed earlier, a way of shaping the minds of officials. Thus differences in assumptions about human nature lay behind their differences about the function of public opinion.

(ii) Elections. Another institution about which Bentham and Rousseau agreed was the value of frequent elections. The electoral incentive was crucial to Bentham's scheme. The essence of creating a responsible legislature was to make the legislator dependent on his constituents but independent of various interest groups, and to instill in the legislator a sense of impermanence. Bentham therefore

recommended annual elections. They would diminish the incentive to act corruptly since the "short-livedness of the power" would "diminish both to producers and thence to sellers the venal value" of office. Secondly, incompetent representative could be got rid of without long delay (PPR, p. 445). Lastly, frequent elections would keep representatives on their toes (PPR, p. 542).

In addition to annual elections, Bentham proposed what he called an "all-comprehensive temporary nonrelocability system" (CC, p. 172). The idea was that no member of the current legislature could run again unless the number of those who had served previously was two or three times the number of those currently in the legislature. This proposal would guarantee that voters had a choice between experienced candidates.

To provide a place for those who could not run again, Bentham proposed the establishment of a "continuation committee," to provide the continuity lacking in a pure system of limited terms. Bentham recognized the importance of balancing the advantages of frequent turnover with those of stability. Since consideration of complex bills might extend over the life of several legislatures, work on them might be unnecessarily interrupted without the continuation committee, which would also help counteract the inexperience of new legislators.

As for Rousseau, though he allowed two-year terms for his legislature, instead of the one-year Bentham recommended, he agreed that "short-livedness" of the legislature was desirable. "England," Rousseau thought, "has lost her liberty for having neglected [to require] frequent re-elections." In England "a single parliament

lasts so long that the court, which would go bankrupt buying it annually, finds it profitable to buy it for seven years, and does not fail to do so" (Poland, p. 188).

Rousseau also recognized the desirability of something not unlike Bentham's "all-comprehensive temporary nonrelocability system," in that he did not allow legislators to succeed themselves indefinitely. He also provided a functional equivalent of the "continuation committee" in that there was to be a constant circulation of the elites between legislature and magistracy.

It is curious that Bentham, who cared much less about continuity (i.e., G4) than Rousseau, should have made such elaborate provision to ensure that there would be carryover from one legislative session to the next, and it is equally odd that Rousseau, who believed that small communities could develop a notion of the common good (G3), should suggest short terms of office for public officials. Given his assumptions about human nature, it is somewhat surprising that he believed that term of office should have a significant effect on legislators' behavior. In a sense therefore, the two theorists agree on the value of short terms of office with some provision for carryover only because both strayed a bit from their assumptions and goals.

V.

#### CONCLUSION

We have proposed a framework in this paper for analyzing proposals that theorists have made for implementing their ideas about

representation. The value of this approach is threefold. First, it provides a standard means of comparing the proposals that theorists make. Often, comparisons are made difficult by the fact that theorists themselves do not use a common language or refer to standard categories so that we tend to see their values, beliefs and proposals as being more idiosyncratic than they actually are. By trying to develop general categories of representation goals and beliefs about human nature, we hope to provide a method for discovering the common choices made in institutional design.

The second value of this framework is that it points out how specific proposals are the joint product of a theorist's goals for representation and his beliefs about human nature. Knowing only how the theorist would like representatives to behave is not sufficient information to understand why he proposed a particular rule or institution: two theorists with the same goals might take different positions on a proposal because their beliefs about human nature differ. Similarly, it is not sufficient to know a theorist's views about human nature. Rather, both kinds of information are necessary conditions.

Finally, this framework helps us to see that theorists are not always consistent with their premises. Sometimes, their proposals betray doubts and uncertainties about their stated beliefs and values. At other times, the connection between a theorist's beliefs and proposals are logically inconsistent. This framework therefore gives us a way of criticizing the specific proposals of political theorists in a systematic and logical manner.

## FOOTNOTES

1. We have used the abbreviated titles in our references for the sake of convenience. The titles for Bentham are as follows:

CC is Constitutional Code

Prom. of the Laws is Essay on the Promulgation of Laws

Intro. to Prin. is An Introduction to the Principles of Morals and Legislation

FG is A Fragment on Government

PPR is Plan of Parliamentary Reform.

All citations of Bentham's writings are from the Bowring edition of his works. The titles for Rousseau are as follows:

Poland is Considerations on the Government of Poland

Emile is The Emile of J.J. Rousseau

Social Contract is The Social Contract

Corsica is Constitutional Project for Corsica

Creed is The Creed of a Priest of Savoy

## REFERENCES

Campbell, Angus et al. The American Voter (New York: Wiley, 1960)

Bentham, Jeremy. The Works of Jeremy Bentham (London: Simpkin, Marshall and Co., 1863), John Bowring ed., 11 vols.

Fenno, Richard. Home Style (Boston: Little, Brown, 1978).

Rousseau, Jean-Jacques. The Creed of a Priest of Savoy trans. by Arthur H. Beattie (New York: Ungar, 1957)

The Emile of J.J. Rousseau trans. and selections by William Boyd (New York: Teachers College Press, 1966)

A Constitutional Project for Corsica in Rousseau: Political Writings trans. by Frederick Watkins (New York: Nelson, 1953)

Considerations on the Government of Poland in Rousseau: Political Writings trans. by Frederick Watkins (New York: Nelson, 1953)

The Social Contract trans. by G.D.H. Cole (New York: Dutton, 1930)

TABLE 1  
THE GOALS MATRIX

	a	b
G1	Short time perspective	Long time perspective
G2	Parochial interests	Common interests
G3	Cooperative ties	Independence
G4	Adaptive to change	Consistent action over time
G5	Freedom of decision	Reliability

TABLE 2  
THE BELIEFS MATRIX

	a	b
B1	Good	Bad
B2	Malleable	Not malleable
B3	Real interests	Apparent interests
B4	Risk averse	Risk acceptant
B5	Secular laws	Transcendental laws