DELIBERATION, CYCLES, AND MISREPRESENTATION

by

Jon Elster

I. Introduction

It is well-known that majority voting can introduce cycles, and that single-peaked preferences provide one sufficient condition under which cycles do not occur. In this paper I shall explore the relation between deliberation and single-peakedness. The main argument is that *publicity* of deliberation (and of the subsequent votes) can reduce the dimensionality of the options and hence reduce the likelihood of cycles. This mechanism may or may not be empirically important; here I shall only argue, by means of a hypothetical example, that public deliberation has at least the potential for preventing cycles.

The idea that deliberation can induce single-peakedness is not new. Dryzek and List (2003) and List et al. (2006) have adduced empirical and theoretical evidence for this connection. Their arguments overlap with mine, notably as regards "the civilizing force of hypocrisy". They do not, however, specify *how* this mechanism might lead to single-peakedness. Although I do not offer a general account, my example will at least provide some insight into how hypocritical misrepresentation of preferences might induce single-peakedness, by forcing certain preference dimensions to go underground, as it were.

In social choice theory, the emphasis is on strategic rather than hypocritical misrepresentation of preferences. The two are related to each other in complicated ways, which I shall try to bring out by two examples: the Borda method for preference aggregation and the passage of the 1964 Civil Rights Act. The second example is particularly striking, since it shows how, under certain circumstances, publicity can have perverse or undesirable effects.

I shall now proceed as follows. In Section **II**, I spell out – with a hefty dose of self-criticism – my current views on the relation between publicity and preference change. To put it briefly, I argue that publicity affects *expressed* preferences rather than sincerely held preferences. In Section **III**, I propose an example – the choice of a property tax level to fund public schools – that shows how a cycle can be made to disappear if debates and votes are public rather than secret. In Section **IV**, I address the issue of strategic misrepresentation and its relation to hypocritical misrepresentation. Section **V** offers a brief conclusion, which to some extent undermines the relevance of the previous Sections.

II. The "civilizing force of hypocrisy"

In any society or community, the members will bestow praise and blame not merely on actions, but also on motivations for acting (For a fuller discussion see Elster 1999, Ch. V). Some motivations are highly ranked, others are more lowly. In classical Athens, they were ranked in roughly the following order: concern for the public interest, anger, self-interest, and envy. In most contemporary Western societies, anger and the desire for vengeance probably occupy a lower place in the hierarchy than the pursuit of self-interest. To cite another motivation to which I shall return in the next Section, *snobbish preferences* are probably also ranked close to the bottom. In the eighteenth-century American elites, there was a cult of disinterestedness (Adair 1998; Wood

1987). Only half a century later, Tocqueville (2004, p. 611) claimed that among the Americans he observed, the pursuit of private interest was ranked above public-interested motivations. This may remain true for contemporary Americans (Miller 1999). In the United States today, therefore, motivations might perhaps be ranked as follows: self-interest, public interest, anger, snobbish preferences, and envy. Assuming that sincere voting is associated with public-spiritedness and strategic voting with self-interested aims, the alleged disapproval in the United States of sincere voting by legislators points in the same direction (Riker 1962, p. 23, as cited by Satterthwaite 1973, p. 13). In societies characterized by "amoral familism" (Banfield 1958), the desire to promote the public interest might be ranked at the very bottom of the hierarchy.

If a person is perceived by others as having acted on a motivation that occupies a low place in the hierarchy, he or she will be exposed to "blame and shame". From Aristotle onwards, many writers have observed that being the target of the contempt of others is an intensely unpleasant experience, independently of any material sanction that might accompany it (Elster 1999, p. 146-47). This fact generates a strong emotionally based second-order motivation for *hypocritical misrepresentation* of first-order motivations. Because others will in general only have a rough-and-ready belief about the *distribution* of the various motivations in the population, such misrepresentation will often succeed. Crucially, the hypocrisy can also, as we shall see, affect non-verbal behavior.

In my earlier work on "the civilizing effect of hypocrisy" I mostly assumed that the desire to promote the public good is at the top of the hierarchy and that individuals will be motivated to present themselves as motivated by that desire. The observations in the paragraph before the last show, however, that the effect of hypocrisy is not always civilizing. In societies with strong codes of honor, even an individual who does not want to take revenge might be forced to do so to avoid

the contempt to which he would be exposed if he didn't. In the United States and elsewhere, public-spirited individuals may fear being branded as do-gooders (Elster 1989, p. 189; see also Hermann, Thöni and Gächter 2008). Amoral familism might also force the general interest to go underground. I shall restrict myself, however, to societies in which the desire to promote the public good is highly ranked, with motives such as malice, envy and snobbishness being close to the bottom. The fear of shame and blame may then keep certain non-public-spirited motivations out of public discussion.

Moreover, and crucially, constraints on public argument may also create constraints on public *action*, notably on public voting. The « consistency constraint » and the « imperfection constraint » can thwart purely opportunistic uses of public-interest arguments (Elster 1999, pp. 347-9, 375-80). First, there is the phenomenon of "moral priming": once an agent has made an opportunistic use of a public-interest argument, she cannot easily abandon it if, on a later occasion, it works against her interest. Second, there is the need to avoid public-interest arguments that coincide *too well* with one's private interest. In practice, the need to respect these two constraints may mimic or simulate true disinterestedness to a considerable degree. *How* considerable is of course an empirical issue.

In a given community, the real or "sincere" motivations can be highly diverse. The civilizing force of hypocrisy may, however, cause *expressed* motivations to be more homogeneous. Since voting cycles can arise from heterogeneous preferences in the population, hypocrisy might, perhaps, alleviate the problem of cycling. In the next Section I propose a hypothetical example – not, in my view, intrinsically implausible - to show how this might happen. First, however, I want to sketch some background.

Many years ago, I wrote that "Rather than aggregating or filtering preferences, the political system should be set up with a view to changing them by public debate or confrontation. The input to the social choice mechanism would then not be the raw, quite possibly selfish or irrational preferences […] but informed and other-regarding preferences. […] There would [then] not be any need for an aggregation mechanism, since a rational discussion would tend to produce unanimous preferences" (Elster 1986, p. 112). To produce this happy outcome, I counted on "the psychological difficulty of expressing other-regarding preferences without coming to acquire them" (Elster 1983, p. 36).

As correctly observed by Christian List et al. (2006), the first statement is hopelessly naive. The authors go on to argue, however, in the general spirit of my argument, that even if discussion does not lead to unanimity, it may make preferences more homogeneous in the sense of moving them towards *single-peakedness* (explained below). Moreover, they adduce empirical evidence that supports this conjecture. Finally, they cite theorems to the effect that the closer preferences are to single-peakedness, the lower the probability of cycles.

The main postulated mechanism in their work is that deliberation makes for more *informed* policy preferences, by changing instrumental and factual beliefs. I shall explore an alternative way in which deliberation followed by a public vote can bring about single-peakedness. If deliberation and voting are public, the civilizing effect of hypocrisy can reduce the heterogeneity of *expressed* preferences and, as a result, prevent cycles from arising.

In the second of the two statements from my earlier work that I cited above, I argued that deliberation might change the sincere or private preferences of the agents. I now make the weaker and, I think, more plausible claim that deliberation followed by public voting might change their *expressed* or public preferences. In my previous work, I misunderstood, in fact, the implications of cognitive

dissonance theory on a crucial point, by assuming that the constraint of having to *state* impartial or public-interest arguments for one's position would always, over time, make one *embrace* those arguments. The theory of cognitive dissonance implies, however, that when an agent expresses an opinion that is contrary to his private view, the strength of the tendency to also adopt the expressed view privately is *inversely related* to the strength of the pressure on the agent to elicit the overt behavior (Festinger and Carlsmith 1959). In a typical experiment, subjects are asked to write an essay on abortion that goes against their own convictions. Subjects who are paid a large sum of money to write the essay do not change their mind, whereas those who are asked to do so as a favor to the experimenter tend to so do. Both types of subject need a *reason* for what they do. For the former, the payment provides a sufficient reason. The latter, however, change their mind to justify their behavior to themselves. Hence dictators cannot rely on dissonance reduction to induce private compliance by forcing public compliance, nor can deliberative democrats count, as I did, on the same mechanism to induce public-spirited *motivations* through the need to use public-spirited *language*. In both cases, the strong *sanctions* that would be triggered were he to express his private opinion provide the agent with a sufficient reason for not doing so.

III. Public schools, private schools, or home schooling?

Imagine now a local community with three groups of voters: senior citizens, young couples without children, and parents with young children. They are, let us assume, equally represented in the municipal council. The personal benefits (utility) each group derives from low, medium and high property taxes and from the corresponding levels of school quality are as shown in Fig.1. Although the options are discrete, the utility functions are drawn as continuous

lines. We may think of the points on the horizontal axis between Low, Medium and High as representing probability combinations of the discrete options.
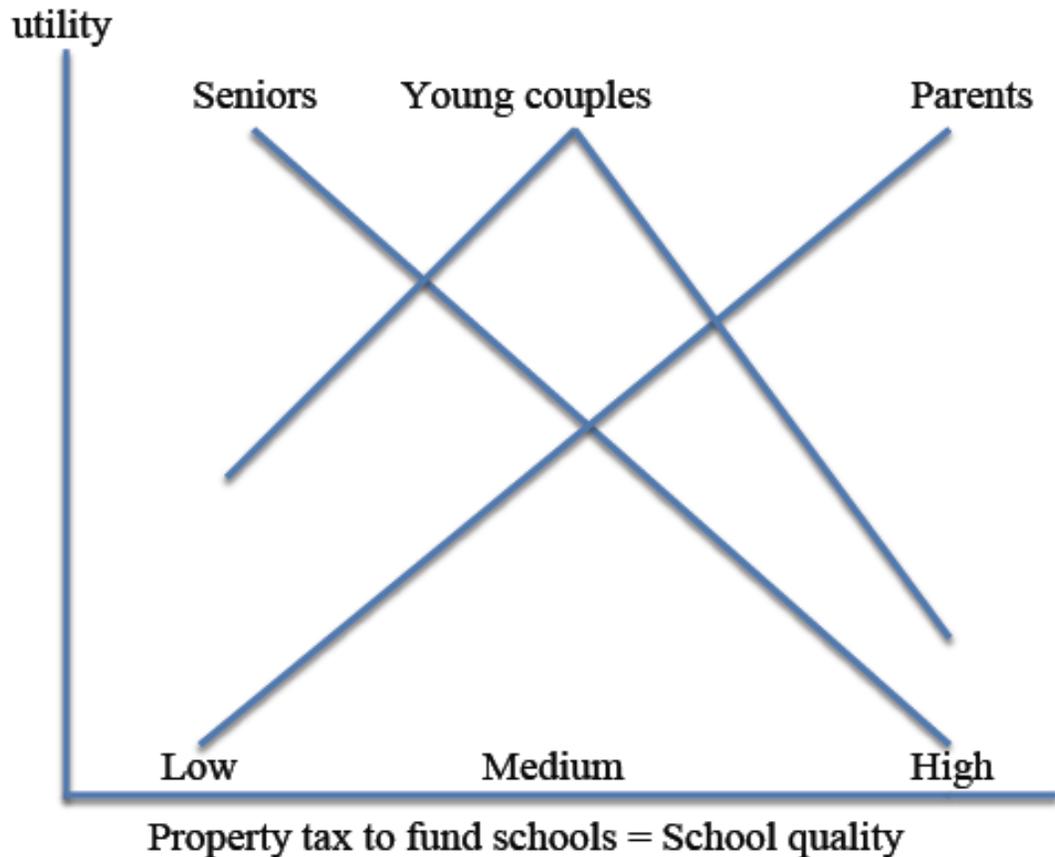

Figure 1

Whereas the utilities for seniors and parents are easily understood, those of the young couples without children require a brief explanation. On the one hand, they have low incomes today. On the other hand, they expect to have children later. In the light of these facts, they prefer a medium-quality school to a bad one and a bad one to an expensive one. If we now assume that the voting preferences of each group are selfish and dictated only by the personal benefits they derive from different levels of school funding, Fig. 1 also represent these preferences. These are

single-peaked, meaning that one can order the options such that for each group the farther to the left or the right an option is from its preferred option (or ideal point) the less it is preferred. In this case, the social preference is unambiguously for medium-level taxes. Young couples and parents prefer this option to low taxes, while young couples and seniors prefer it to high taxes.

Assume now a different scenario. Instead of the seniors, there is a group of "snobs". Since they want to place their children in a private school (assumed to be neither better nor worse than a high-quality public school), their top-ranked option is to pay low taxes that will allow them to pay the high tuition fees. At the same time, their bottom-ranked option is a medium tax level that will neither allow them to pay the tuition fees nor ensure for their children a high-quality education. Fig. 2 shows the ensuing private preference profile, which is not single-peaked. If people vote following their private preferences, as they would do in a secret ballot, majority voting yields a cycle.
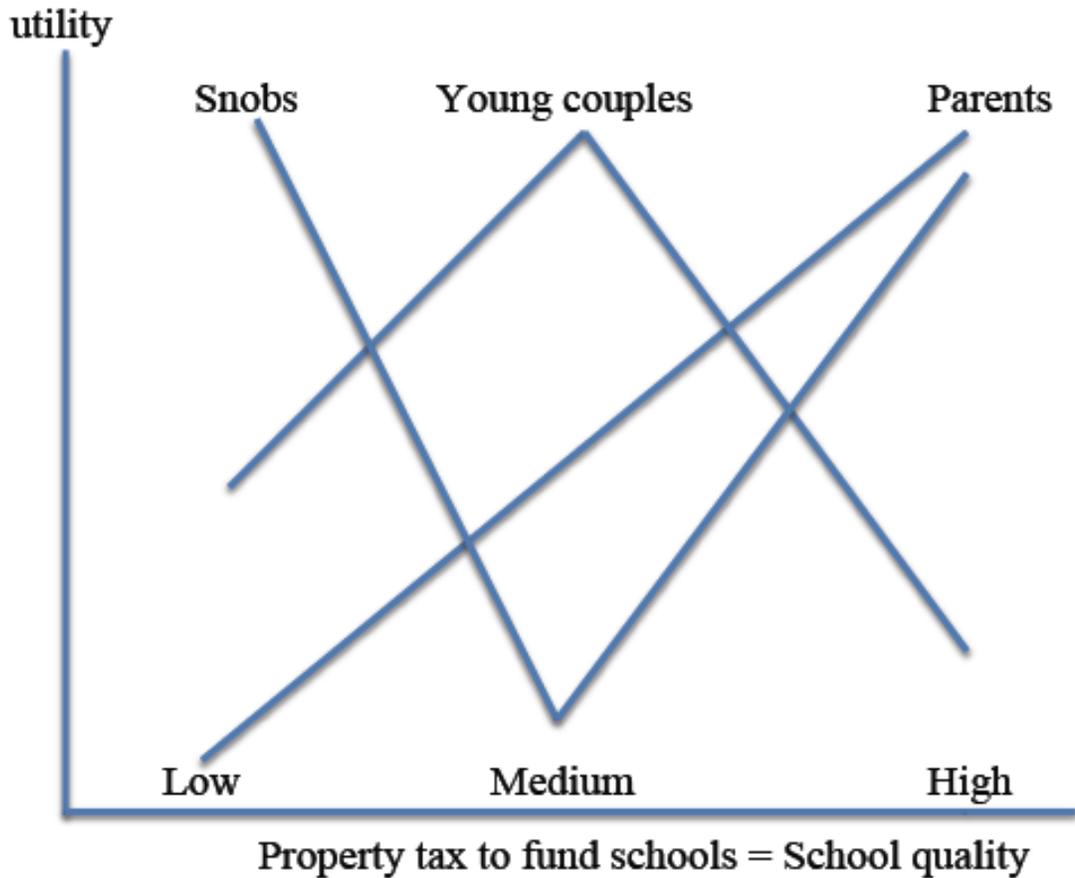
Snobs and young prefer L to H. Snobs and parents prefer H to M. Parents and young prefer M to L. Hence *interest-based secret majority voting* yields a cycle.
Fig. 2

My suggestion is that if deliberation and voting are public, the snobs will be *ashamed to express their snobbish private preferences*. They would not be willing to state publicly that they do not want their children to go to the same (high-quality) public school as children from low-income families. Instead, they would express preferences identical to those of the parents. This would illustrate the civilizing force of hypocrisy.

To recapitulate, I have offered three hypothetical examples of majority voting with three voting blocs and three options. (i) In the senior-young-parent case, private preferences are single-peaked. Under a secret ballot in which the blocks vote on the basis of these preferences, no cycle arises. It seems plausible that the same result would be observed under public voting. Although there often attaches an opprobrium to self-interested voting, I believe that in the kind of situation I have described, voters would see it as perfectly acceptable if others vote according to their self-interest, since it is not at all clear what the public interest would require. (ii) In the snob-young-parent case with secret ballot, expressed preferences are not single peaked, because the snobs now introduce an additional dimension to the problem. (iii) In the snob-young-parent case with public voting, this dimension goes underground, and the expressed preferences are single-peaked. In case (i) as well as in case (iii), expressed preferences are relatively homogeneous. There is conflict, but it is manageable. In case (ii), it is unmanageable because of the greater heterogeneity of expressed preferences.

I do not imply, of course, that *all* multi-dimensional conflicts involve "inadmissible" or "unavowable" preferences along all dimensions except one, and that publicity of debates and votes would *always* eliminate cycling. The preferences of the snobs as shown in Fig. 2 could also be due to motives that occupy a higher place in the normative hierarchy. Suppose, for example, that some parents believe that home-schooling provides the best education, although it would require one parent to stay at home without earning an income. They could afford and would prefer this option with low school taxes, but could not afford it with medium or high taxes. Yet because they place a high priority on a good education for their children, they would prefer the best public education to a mediocre one. In other words, the dimension of private versus public schools is here replaced by that of home versus public schooling. Unlike the views of the snobs, the preference of

the home-schoolers can perfectly well tolerate the daylight of publicity. Public deliberation and voting would not suppress the cycle.

It is clearly unfeasible to assess how *frequently* this cycle-preventing effect of public deliberation occurs. Perhaps the phenomenon is empirically marginal. I believe nevertheless that it merits a place in the catalogue of cycle-preventing conditions and mechanisms. The civilizing force of hypocrisy can reduce the dimensionality of *overt* conflict, and cycling often arises from multi-dimensionality.

IV. Strategic and hypocritical misrepresentation

In collective decision-making, strategic behavior can take a number of forms.

First, there is strategic misrepresentation of preferences. A person may choose to *express* preferences that differ from her sincere preferences, in order to increase the likelihood that the option she sincerely prefers will be chosen. The Gibbard-Satterthwaite theorem states that all deterministic voting systems are manipulable in this sense.

Second, there is the strategic introduction of additional options, as in the "killer amendments" further discussed below.

Third, there is the strategic choice of procedure. In a first example, assume that there are three groups in the House of Representatives of roughly equal size I, II and III, which rank the options as follows, inducing a majority cycle:

I	II	III
B	A	C
A	C	B
C	B	A

Fig. 3

If the leaders of group I can somehow persuade the others to vote by holding the winner in a vote between A and C up against B, their preferred alternative will be chosen. Below I consider cases when the order of pairwise voting between alternatives is fixed rather than manipulable.

For another example, involving aggregation of beliefs rather than of preferences, consider a "doctrinal paradox":

	Did the accused do X?	Is doing X breaking the law?	Is X guilty?
Judge 1	Yes	Yes	Yes
Judge 2	Yes	No	No
Judge 3	No	Yes	No

Table 1

If the judges follow what is perhaps the standard procedure, they will vote on the verdict as shown in Table 1 (*aggregation of conclusions*) and the accused will

be acquitted. If, however, they first decide on the facts by majority voting and then on the law by majority voting (*aggregation of premises*), the accused will be convicted.

Suppose that the court uses aggregation from premises. Then either Judge 2 or Judge 3 could move the decision of the court to the outcome they prefer by misrepresenting their opinions on, respectively, the factual and the legal issue. Moreover, the judges could have an incentive to manipulate the procedure itself. Experimental evidence suggests that when a group has to choose between aggregation of premises and aggregation of conclusions, the choice is essentially made as a function of the expected outcome (Bonnefon 2007, 2010). Generalizing this idea to the legal case, suppose that only Judge 1 knows the views of the other judges. To achieve his favored verdict of Yes, he might persuade a majority of the court to adopt the premise-based procedure. If, however, Judge 2 were able to anticipate this move, he could defeat it by changing his opinion on the factual premise to No. Hence a procedural manipulation might trigger the misrepresentation of belief.

In the Borda count, each participant lists the options in her order of preference. The election official then assigns the number 1 to her first-ranked option, the number 2 to the second-ranked and so on, adds up for each option the numbers assigned to it, and declares as the winner the option whose numbers sum to the smallest total. Normatively, this procedure has many attractive properties (Mackie 2003, Ch.3). In particular, it captures an important *consensual* aspect of collective decision-making by preventing the choice of candidates that are top-ranked by a bare majority and bottom-ranked by a large minority. In any given case, this fact may be irrelevant, for two reasons. First, the ordinal ranking of each voter does not provide information about the cardinal intensity of preferences. Second, even if it did, it would not allow us to compare interpersonal strength of

preferences. I conjecture (but cannot prove) that in the long run and on average, the Borda count would nevertheless enhance the consensual nature of the decision and reduce the potential for social conflict.

The Borda count does, however, generate an incentive for misrepresenting preferences. Assume that **x** would win by one point under sincere preferences, but a voter that has **y** as her first-ranked candidate misrepresents her preference ordering y > x > z > v > v as y > z > v > v > x. The manipulation will enable her to get **y** elected, assuming that all others vote sincerely. Others may, however, respond in kind. As explained by Satterthwaite (1973, Ch. 1.2), this behavior can have a number of undesirable consequences.

In some cases, *publicity* may provide a remedy. Even a strategic agent may be embarrassed to "lie in public" by expressing preferences that everybody knows she does not hold. In this case, others would sanction *the insincerity of the expressed preferences*, whereas in the pure hypocritical case they would sanction *the substance of the private preferences* were these to be expressed. In the case of the Borda count, the fact of ranking one's second choice in the last place would also reveal that one *cared less about making a good collective decision than about having one's own candidate chosen*. When someone pointed out to Borda that his scheme lent itself to manipulation, he is reported to have answered that "My system is intended only for honest men", implying that attempts to game the system would reveal dishonesty. It is possible, therefore, that publicity might be an antidote not only to cycling, but also to strategic voting. To put it more dramatically: Bentham - the great advocate of publicity in political theory - might take some of the sting out Arrow's impossibility theorem as well as of the Gibbard-Satterthwaite theorem. How much of the sting remains is an open question.

In other cases, publicity may have bad effects. Consider Fig. 3 again. Assume first that A is the status quo and that B is proposed as an alternative. In a pairwise vote, groups I and III will ensure that B is chosen. Assume, however, that group II strategically proposes C as a "killer amendment", inducing a cycle. Under the rules of Congress, C will first be held up against B. Assuming sincere voting, C will be the winner. Next, C will be held up against A. Assuming sincere voting, the status quo will be the winner. Anticipating this outcome, however, members of I and II may vote strategically against the killer amendment, to ensure the adoption of the original bill. Strategic introduction of a new option may induce strategic misrepresentation of preferences.

The following story shows that in the passage of the Civil Rights Act of 1964, publicity almost caused the adoption of a killer amendment:

> [When] the 1964 Civil Rights Act was being debated in the House, the conservative chair of the Rules Committee (Howard Smith of Virginia) introduced what he thought of as a killer amendment: he proposed that the bill prohibit discrimination on grounds of sex as well as race and national origin. The House leadership (sharing Smith's belief that the amendment might kill the bill) tried to persuade liberal Democrats to vote against the proposal. According to some accounts, many of these members would have been personally willing to do so if only such a vote could have been held in private. But the galleries were filled with women's interest groups who were there to observe behavior on any unrecorded (teller) votes and, at least in this case, liberal members were unwilling to go along with the leadership and vote down Smith's amendment. As it turns out the Smith amendment did not lead to the defeat of the bill but rather to the inclusion of gender as one of a group of statutorily protected classes. We suspect that liberal congressmen

were favorably disposed to prohibiting gender discrimination all along but that they would have preferred to follow the Democratic leadership if such work could have been done in the dark of night. That is in a secret vote, many would have had an induced preference to vote against the Smith amendment. However, the fact that the vote was to be "public" (in the sense of a standing or teller vote) and the presence of motivated interest group representatives probably induced these members to vote for the amendment (Brady, Ferejohn and Pope 2007, p. 70).

Since the liberal democrats thought that the House leaders and Smith were correct in assuming that the amendment would kill the Civil Rights Act, they would have been justified in voting against their conviction. If the assumption had actually been right, exposure to the public of feminists would have prevented the passage of the Civil Rights Act: the best would have been the enemy of the good. As it turned out, the assumption was wrong and the amended Act was passed. Publicity had a benign effect, but only by accident.

This case is unusual in involving three layers of preferences: the sincere preferences of the liberal Democrats for the amendment, their strategic motivation to misrepresent these preferences, and their hypocritical motivation to express their sincere preferences. The example shows that publicity is not a panacea, that strategic misrepresentation of preferences is not always undesirable, and that hypocritical misrepresentation induced by a public-spirited audience is not always desirable.

V. Conclusion

Although cycles, misrepresentation and killer amendments have attracted much theoretical attention, their empirical importance is questionable. Mackie

(2003) has successfully challenged every single instance of the alleged voting cycles adduced by William Riker, as well as virtually all other putative cycles discussed in the social-choice literature. While Mackie does not claim that cycles never occur, he believes they are marginal. In an unusually thoughtful paper, Johansen (1977) argued that the issue of misrepresentation of preferences, specifically in the provision of public goods, is highly artificial. As for killer amendments, Finocchiaro and Jenkins (2008) only uncovered five successful instances in the US Congress between 1953 and 2004.

I shall not address the (to me) puzzling paucity of demonstrated cycles. It seems clear (to me), however, that strategic voting as well as strategic introduction of new alternatives are unlikely to occur frequently in the intensely interactive political arena. Having noted the lack of empirical evidence for misrepresentation, Johansen (1977, p. 148) goes on to say that

> [One] might argue that it is implied by the nature of the issue that historical records and empirical evidence would never be available, if the persons and parties involved are efficient and consistent in misrepresenting their preferences. I do not, however, consider this to be a convincing argument. If the theory of concealment of preferences were correct, then practically everyone involved in public affairs would be a practitioner of the art, and he would probably understand that everyone, foe and friend alike, was practicing the same tricks. Add to this the fact that many people like to reveal misbehavior of their adversaries and boast of tricks performed by themselves if they have been successful. Then it is hard to believe that we should not get ample empirical evidence if concealment of preferences were as important in practice as it seems to be in theory.

Along similar lines, Mackie (1998) also argues against the idea that "All men are liars". The objections adduced by him and by Johansen do not amount to a claim that strategic misrepresentation of preferences and other strategic moves *never* occur. They do, however, suggest that this behavior is not at the core of politics. The objections may not, perhaps, carry over to *hypocritical* misrepresentation. Because of the imperfection constraint and the consistency constraint, hypocritical behavior can never be as relentlessly opportunistic as, according to the theory, strategic behavior is supposed to be.

REFERENCES

Adair, D. (1998), *Fame and the Founding Fathers*, Indianapolis: Liberty Press

Banfield, E. (1958), *The Moral Basis of a Backward Society*, Glencoe, Ill.: The Free Press.

Bonnefon, J.-F. (2007), « How do individuals solve the doctrinal paradox in collective decisions », *Psychological Science* 18, 753-55.

Bonnefon, J.-F. (2010), "Behavioral evidence for framing effects in the resolution of the doctrinal paradox », *Social Choice and Welfare* 314, 631-41.

Brady, D., Ferejohn, J. and Pope, J. (2007), "Congress and civil rights policy: An examination of endogenous preferences", in I. Katznelson and B. Weingast (eds.), *Preferences and Situations*, New York: Russell Sage, pp. 62-86.

Dryzek, J. and List, C. (2003), "Social choice theory and deliberative democracy: A reconciliation", *British Journal of Political Science* 33, 1-28.

Elster, J. (1983), *Sour Grapes*, Cambridge University Press.

Elster, J. (1986), "The market and the forum", in J. Elster and A. Hylland (eds.), *Foundations of Social Choice Theory*, Cambridge University Press, pp. 103-32.

Elster, J. (1989), *The Cement of Society*, Cambridge University Press

Elster, J. (1999), *Alchemies of the Mind*, Cambridge University Press

Festinger, L. and Carlsmith, J. (1959), "Cognitive consequences of forced compliance", *Journal of Abnormal and Social Psychology* 58, 203-10.

Finocchiaro, C. and Jenkins, J. (2008). "In search of killer amendments In the modern U.S. House", *Legislative Studies Quarterly* 33, 2`63-94.

Hermann, B., Thöni, C. and Gächter, S. (2008), "Antisocial punishment across societies", *Science* 319, 1362-67.

Johansen, L. (1977), "The theory of public goods: Misplaced emphasis?", *Journal of Public Economics* 7, 147-52.

List, C. et al (2006), "Deliberation, single-peakedness and the possibility of deliberative democracy", Working Paper, London School of Economics.

Mackie, G. (1998), "All men are liars: Is democracy meaningless?", in J. Elster (ed.), Deliberative Democracy, Cambridge University Press, pp. 69-96.

Mackie, G. (2003), *Democracy Defended*, Cambridge University Press.

Miller, D. (1999), "The norm of self-interest", *American Psychologist* 54, 1053-60.

Riker, W. (1962), *The Theory of Political Coalitions*, New Haven CO: Yale University Press.

Satterthwaite, M. (1973), *The Existence of a Strategy Proof Voting Procedure*, Ph.D. Dissertation, University of Wisconsin.

Tocqueville, A. de (2004), Democracy in America, New York: *Library of America*.

Wood, G. (1987), "Interest and disinterestedness in the making of the constitution", in R.Beeman, S.Botein and E. Carter (eds.), *Beyond Confederation*, Chapel Hill: North Carolina Press 1987, pp.69-109.