CHAPTER 6

IS DEMOCRACY MEANINGLESS? ARROW’S CONDITION OF THE INDEPENDENCE OF IRRELEVANT ALTERNATIVES

Introduction. The interpretation that Arrow’s Condition I, independence of irrelevant alternatives, prohibits the use of individuals’ intensities of preference in the construction of social choices, is not precise. Rather, it is the social welfare function (as defined in Chapter 4), which demands both individual and social orderings, and thereby prohibits cardinal utility inputs. Condition I, as Arrow wrote it, redundantly requires individual orderings, but goes further and demands that, even given the ordinal data from individual orderings, the social choice over any two alternatives not be influenced by individuals’ preferences involving any third alternatives. This is explicit in Arrow (1963/1951, 59, emphasis added):

It is required that the social ordering be formed from individual orderings and that the social decision between two alternatives be independent of the desires of individuals involving any alternatives other than the given two . . . . These conditions taken together serve to exclude interpersonal comparison of social utility either by some form of direct measurement or by comparison with other alternative social states . . .

Arrow’s is a strong independence condition. Slight weakenings of it allow the Borda count or the Young-Kemeny rule as possible, that is, nondictatorial, social welfare functions, and further weakenings permit further voting procedures.
Barry and Hardin (1982, 217-218) agree that Arrow’s IIA is a powerful condition. “Part of its power is that one cannot easily intuit what it means or why it matters. . . . Perhaps because of its subtlety, condition I is apparently the condition that is most readily taken for granted in the proof of Arrow’s and related theorems.” Its content is frequently misunderstood. Justifications of the condition are typically thin and dogmatic, often no more than an assertion that its appeal is intuitively obvious. My search for justifications of the condition found thicker arguments mostly by Arrow (1963/1951, 1952, 1969, 1987, 1997), Sen (1970, 1982), some by Riker (1961, 1965, 1982), and otherwise mostly repetition of points made by Arrow without further justificatory development.3

The chapter proceeds as follows. First, I explain that historically many people have misunderstood the content of the independence condition (IIA(A)), believing it to be another condition (IIA(RM)), one that does not contribute to the impossibility result. My main point is that the independence condition can not be defended as intuitively obvious if sophisticated commentators have trouble grasping even its content. Second, I show by example that to violate either Arrow’s independence condition (IIA(A)) or the contraction-consistency independence condition (IIA(RM)) can be substantively rational. I point out that Arrow understands the simplifying assumptions of his model not as ends in themselves but as means to empirical analysis; the assumptions themselves have no descriptive or normative force, even more so when they are contrary to observation and intuition. Third, Arrovians defend the IIA(A) as forbidding the influence of irrelevant alternatives over the consideration of relevant alternatives in social choice. I explain that the condition has nothing to do with forbidding consideration of irrelevant alternatives in the ordinary sense of the term, but rather requires that social choice be carried out only by pairwise
comparison, thereby delivering the impossibility result. Fourth, I continue discussion of the Arrow theorem as motivated to avoid interpersonal comparisons of utility, and argue that the IIA(A) is superfluous to that goal. I submit that the Borda and Condorcet methods are equivalent with respect to comparing or not comparing mental states, and that both are purely ordinal methods. Fifth, I examine Riker’s claims that the IIA(A) serves to forbid undesirable utilitarian voting rules, probabilistic voting rules, and consideration of irrelevant alternatives. I reply that it has not been shown that utilitarian or probabilistic voting rules are undesirable, but if they are then there are weaker independence conditions that forbid those voting rules but still permit other rules such as the Borda count. Also, I point out that the concern over irrelevant alternatives may be of importance in social welfare applications, but is of no importance in most voting applications. Sixth, I challenge the Arrovian view that the Borda count is logically susceptible to manipulation by addition and deletion of alternatives but that the Condorcet method is not. I argue that both are logically susceptible to such manipulation, but that the susceptibility is of little practical importance. I conclude that the frailties of the reasonable voting rules have been much exaggerated, and that the time has come to move from a destructive constitutional “physics” to a constructive constitutional “engineering.”

The Wrong Principle is Defended. Arrow is an economist, and analogizes political choice to economic choice. He explicitly considers voting and the market as special cases of the more general category of social choice. Arrow borrows his basic conception from simple consumer economics. He distinguishes all possible or conceivable alternatives from all feasible or available alternatives. Suppose that the set $X$ contains all possible or conceivable alternatives. $S$, a nonempty subset of $X,$
contains all feasible or available alternatives. To anticipate, \( S \) contains the “relevant” alternatives and \( X - S \) contains the “irrelevant” alternatives.

On any given occasion, the chooser has available to him a subset \( S \) of all possible alternatives \( [X] \), and he is required to choose one out of this set \( [S] \). The set \( S \) is a generalization of the well-known opportunity curve; thus, in the theory of consumer’s choice under perfect competition it would be the budget plane. It is assumed further that the choice is made in this way: Before knowing the set \( S \), the chooser considers in turn all possible pairs of alternatives, say \( x \) and \( y \), and for each such pair he makes one and only one of three decisions: \( x \) is preferred to \( y \), \( x \) is indifferent to \( y \), or \( y \) is preferred to \( x \) . . . . Having this ordering of all possible alternatives, the chooser is now confronted with a particular opportunity set \( S \). If there is one alternative in \( S \) which is preferred to all others in \( S \), the chooser selects that one alternative. (Arrow 1963/1951, 12).

This will be a crucial passage.

Many people, including myself for many years, have misunderstood the content of Arrow’s independence condition. Indeed, about 20 years after first publication of his theorem it was recognized that Arrow in 1951 at one point seemed narratively to justify a condition that was not the same as the formally stated condition necessary for his proof. 4

Again on analogy to consumer choice, Arrow (1963/1951) argues that a social choice from a set of alternatives \( S \), “just as for a single individual,” should be
independent of alternatives outside of $S$. He illustrates with a criticism of the Borda count:

For example, suppose . . . an election system . . . whereby each individual lists all the candidates in order of his preference and then, by a preassigned procedure, the winning candidate is derived from these lists. . . . Suppose that an election is held, with a certain number of candidates in the field . . . and then one of the candidates dies. Surely the social choice should be made by taking each of the individual’s preference lists, blotting out completely the dead candidate’s name, and considering only the orderings of the remaining names in going through the procedure of determining the winner. That is, the choice to be made among the set $S$ of surviving candidates should be independent of the preferences of individuals for candidates not in $S$. To assume otherwise would be to make the result of the election dependent on the obviously accidental circumstance of whether a candidate died before or after the date of polling. Therefore, we may require of our social welfare function that the choice made by society from a given environment depend only on the orderings of individuals among the alternatives in that environment. (Arrow 1963/1951, 26)

Arrow’s “surely” is too quick. The election has already provided us with a social ranking. Rather than deleting the dead candidate’s name from each individual’s preference list, why not instead delete the dead candidate’s name from the social ranking? Arrow’s example is thus: two voters rank the alternatives $x > y > z > w$, and one voter ranks them $z > w > x > y$. The Borda ranking is $x > z > y > w$, Arrow’s
focus is that $x$ wins (the Condorcet ranking is $x > y > z > w$). Now candidate $y$ is deleted. By the Condorcet method the ranking of remaining alternatives stays the same, $x > z > w$, and $x$ would still be the winner. If the Borda method is reapplied to the remaining three candidates, however, then the Borda ranking changes to $(x \sim z) > w$, and both $x$ and $z$ are tied for the win.

Then, certainly, if $y$ is deleted from the ranks of the candidates, the system applied to the remaining candidates should yield the same result, especially since, in this case, $y$ is inferior to $x$ according to the tastes of every individual; but, if $y$ is in fact deleted, the indicated electoral system would yield a tie between $x$ and $z$. Arrow (1963/1951, 27)

There are two problems. First, recall that in Arrow’s scheme the chooser, before knowing the set $S$, forms a ranking over all possible alternatives in $X$. Then the chooser encounters $S$, the set of all feasible alternatives. The chooser consults the list made from $X$ and selects the highest ranking alternative or alternatives in $S$ as the choice. Arrow analogizes social choice to the economic model of individual choice. If the analogy holds, then social choice should form a ranking over all possible alternatives $X$, and consult from the list made from $X$ in order to decide the winner or winners in $S$. As this would apply to Arrow’s story about the Borda count and the dead candidate, the Borda count would be carried out on the set $X$ of four candidates and the ranking $x > z > y > w$ determined. Then $y$ dies. We do not reapply the Borda count to the three remaining candidates in set $S$, rather we consult the ranking over four candidates, $x > z > y > w$, and delete $y$ from the social ranking, for $x > z > w$, and $x$ remains the winner. If we take the Borda count over $X$, call that the *global* Borda rule, and if we take the Borda count over some subset $S$ call that the *local* Borda rule.
The general idea of Arrow’s scheme suggests that we should apply the global Borda rule to $X$, but in this instance Arrow says we should apply the local Borda rule to $S$. Why the inconsistency?

The second and much bigger problem is that Arrow’s example does not illustrate the IIA condition used in the theorem, but rather a different condition confusingly labeled “independence of irrelevant alternatives” by Radner and Marschak (1954), also called by Sen “Condition α,” and also called “contraction consistency.”

IIA(Arrow): Let $R_1, \ldots, R_n$ and $R'_1, \ldots, R'_n$ be two sets of individual orderings and let $C(S)$ and $C'(S)$ be the corresponding social choice functions. If, for all individuals $i$ and all $x$ and $y$ in a given environment $S$, $xR_iy$ if and only if $xR'_iy$, then $C(S)$ and $C'(S)$ are the same. (Arrow 1963/1951, 27).

IIA(Radner-Marschak): If $x$ is an element of the choice set of $S$ and belongs to $S_i$ contained in $S$, then $x$ is also an element of the choice set of $S_i$, i.e., $x \in C(S)$ and $x \in S_i \subset S$ together imply $x \in C(S_i)$ (Ray 1973, 987, after Radner and Marschak).

The two conditions are logically independent of one another (Ray 1973). In 1951, Arrow (1963/1951, 32-33) criticized by example the idea that summation of normalized von Neumann-Morgenstern cardinal utility functions might serve as a social welfare function. His first objection was an example that violated what is here called IIA(RM) or contraction consistency, and his second objection was an example that violated the IIA(A). In this passage, Arrow distinguishes the two conditions, but it seems that he does not in the passage justifying the IIA(A) by appeal to the example.
of the dead-candidate in the Borda count (1963/1951, 27). Bordes and Tideman (1991) provide ingenious constructions that would make Arrow both accurate and consistent in his justification of the IIA(A), and their interpretation is more than merely plausible. In the end, I do not join in Bordes and Tideman’s charitable reading, simply because Arrow (1987, 195) later acknowledged a mixup: Nash’s condition (adapted by Radner and Marschak), he said, “refers to variations in the set of opportunities, mine to variations in the preference orderings. . . . The two uses are easy to confuse (I did myself in Social Choice and Individual Values at one point).”

The later Arrow (1997) explains that the key conditions of the theorem are Collective Rationality and IIA(A). For a given election the problems posed by the conditions are hypothetical or counterfactual, he says. First, what would have happened if we had added a candidate who wouldn’t have won or if we had subtracted a losing candidate? The addition or subtraction of such “irrelevant” candidates could have changed the outcome of the election, and this would be a violation of the Collective Rationality Condition, according to Arrow. The Collective Rationality Condition requires that “for any given set of [individual] orderings, the social choice function is derivable from an [social] ordering” (Arrow 1969, 70), which in turn requires IIA(RM). Second, what would have happened if voter’s preferences over noncandidates changed? Change of preference over “irrelevant” candidates could have changed the outcome of the election, and this would be a violation of the IIA(A) condition, according to Arrow. What “this argues is that the election rule was such that the result actually obtained might have been different although it should not have been” (1997, 5). Notice that Arrow’s conditions require that the social choice not change if individual preferences or availability of alternatives were to change. A more natural expectation would be that social choice may or may not change if
individual preferences or availability of alternatives were to change. Consider simple
majority rule over two alternatives – if one member of the majority changes her vote
to the minority position, that may or may not change the social choice. It seems to me
that, in response to changes in individual preferences or in availability of alternatives,
any demand that the social choice should always change, or, the demand by the
IIA(A) or IIAR(M), that the social choice should never change, carries the burden of
justification.

I will now illustrate violation of IIA(A). Suppose that there are two voters
who rank $A > B > C$, two who rank $B > C > A$, and one who ranks $C > A > B$. By the
Condorcet method the collective outcome from the profile is the cycle $A > B > C > A$
and by the global Borda method the collective outcome is $B > A > C$. In order to
investigate violation of the IIA(A) we are interested only in the rankings of two
alternatives, say alternatives $A$ and $B$. Suppose now, counter to the first supposition,
that the two who rank $B > C > A$, instead rank $C > B > A$. By the Condorcet method
the collective outcome changes from a cycle to $C > A > B$, and by the global Borda
method from $B > A > C$ to $C > A > B$. Focus on the Borda count. Under the first
supposition, the Borda count yields $B > A$. Voters’ preferences over the pair $A$ and $B$
do not change, but two voters change from ranking $C$ second to ranking $C$ first. Then,
under the second supposition, the Borda count yields $A > B$, a reversal from the first
supposition. The IIA(A) is violated.

Table 1. Violation of IIA(A)

<table>
<thead>
<tr>
<th>IIA(A)</th>
<th>Actual</th>
<th>Counterfactual</th>
</tr>
</thead>
<tbody>
<tr>
<td># voters: Rank:</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1st</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>2nd</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>3rd</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>Cond.</td>
<td>$A &gt; B &gt; C &gt; A$</td>
<td>$C &gt; A &gt; B$</td>
</tr>
</tbody>
</table>
I will now illustrate violation of IIA(RM), or of contraction consistency.

Suppose again that there are two voters who rank $A > B > C$, two who rank $B > C > A$, and one who ranks $C > A > B$. Again, by the Condorcet method the collective outcome from the profile is the cycle $A > B > C > A$ and by the local Borda method the collective outcome is $B > A > C$. In order to investigate violation of the IIA(RM) we are interested only in the rankings of two alternatives, say alternatives $A$ and $B$.

Suppose now, counter to the first supposition, that instead of the three alternatives $A$, $B$, and $C$, there are only two alternatives, $A$ and $B$. By the Condorcet method the collective outcome changes from $A > B > C > A$ to $A > B$, and by the global Borda method from $B > A > C$ to $A > B$. Again, focus on the Borda count. Under the first supposition, the Borda count yields $B > A$. Voters’ preferences over the pair $A$ and $B$ do not change, but alternative $C$ is removed from the menu. Thus, under the second supposition, the Borda count yields $A > B$. $B$ is an element of $\{A, B\} \subset \{A, B, C\}$, and $B$ wins among $\{A, B, C\}$, but does not win between $\{A, B\}$ – contraction consistency says that if $B$ wins $\{A, B, C\}$ then it should win $\{A, B\}$. The IIA(RM) is violated.

<table>
<thead>
<tr>
<th>IIA(A)</th>
<th>Actual</th>
<th>Counterfactual</th>
</tr>
</thead>
<tbody>
<tr>
<td># voters:</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Rank:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>2nd</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>3rd</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>Cond.</td>
<td>$A &gt; B &gt; C &gt; A$</td>
<td>$A &gt; B$</td>
</tr>
<tr>
<td>Borda</td>
<td>$B &gt; A &gt; C$</td>
<td>$A &gt; B$</td>
</tr>
</tbody>
</table>
In Arrow’s 1951 presentation (1963/1951), an axiom was presented that the individual preference relation is complete, and another presented that it is transitive, and it was further stated that a relation that satisfied those axioms was a weak ordering (12). Individual orderings and social orderings must satisfy the two axioms (19). A social welfare function is a rule which for each set of individual orderings of alternatives states a corresponding social ordering of alternatives (23). A social choice function $C(S)$ is the set of alternatives $x$ in $S$ such that, for every $y$ in $S$, $x$ is weakly preferred to $y$. A social welfare function determines a unique social choice function, and the social choice function satisfies contraction consistency (Bordes and Tideman 1991, 170). A voting rule that violates contraction consistency violates the requirements of the Arrow theorem. Hence, Arrow’s scheme does require that both IIA(RM) and IIA(A) be satisfied. Sen (1993) proved that an impossibility result can be reached even if the requirement for contraction consistency, IIA(RM), is dropped; a condition similar to IIA(A) must be retained, however. IIA(A) is the culprit in the impossibility result.

Arrow’s possibility theorem shows that social ordering, universal domain, pareto principle, nondictatorship and IIA(A) are inconsistent. Ray (1973) shows that social ordering, universal domain, pareto principle, nondictatorship and IIA(RM) are consistent. All social welfare functions satisfy IIA(RM), and at least one, the Borda method, satisfies the remaining conditions, according to Ray. All social welfare functions satisfy IIA(RM) if, as Arrow originally suggested, they are taken on $X$: if we apply our voting rule to all possible alternatives and consult that ranking to order any subset $S$ of $X$, then no contraction of the set of alternatives has taken place. If it is insisted that the voting rule be applied to $X$, and then be reapplied to $S$, then violations of contraction consistency are possible.
The *global* Borda count, for example, taken on $X$, satisfies IIA(RM), contraction consistency, but may violate Arrow’s IIA(A). It satisfies contraction consistency because with the global Borda count there is no contraction from $X$ to $S$. The global Borda count may violate Arrow’s IIA(A) as shown in the example in Table 1. The *local* Borda count, taken on $S$, satisfies Arrow’s IIA(A), but may violate IIA(RM). To determine the rankings of $A$ and $B$, the local Borda count is applied only to $A$ and $B$, $C$ does not enter the picture, thus the local Borda count outcome is the same before and after the change by two voters from $B > C > A$ to $C > B > A$ – and there is no violation of Arrow’s IIA(A). The local Borda count may violate IIA(RM) as shown in the example in Table 2, where, contracting from three alternatives to two alternatives, the outcome changes from $B$ to $A$.

There is a wide discourse, illustrated in my hall of quotations, declaring that democracy is dubious because of the Arrow theorem. The problem I am getting at here is that *many* commentators have made this claim on the mistaken view that the Arrow theorem depends on the IIA(RM) rather than on the IIA(A) (see, e.g., Riker 1961, Riker 1965). Some of these commentators spend time justifying IIA(RM), believing they are thereby justifying the impossibility result. But IIA(RM) is not essential to the impossibility result, rather it is IIA(A) that is essential, and IIA(RM) does not lead to impossibility. Would not such a discovery suggest a revision in views? Instead, what we see in some commentators (Riker 1982) is not a revision in view, but rather a new attempt to justify the newly understood IIA(A). The conclusion is driving the premises, the tail is wagging the dog.

I have made many astounding errors in earlier drafts of this volume and I fear that, despite my best efforts, astounding mistakes and misunderstandings remain. Scholars more talented and diligent than I are bound to err, because the issues under
consideration are difficult. My purpose here is not to dwell on the errors of others. The purpose is to call into doubt the common assertion that the IIA(A) condition should be accepted just because it is intuitively obvious. How could the condition be intuitively obvious if many sophisticated commentators are confused even about its content, let alone its implications?

**The Independence Conditions Are Not Always Substantively Rational.**

Barry and Hardin (1982, 266) say that, “Nobody has any immediate views about the desirability of, say, the independence of irrelevant alternatives, and we should refuse to be bullied by a priori arguments to the effect that we would be ‘irrational’ not to accept it.” Arrovians proceed as if IIA(RM) and IIA(A) were requirements of rationality. One or the other of the conditions is presented as intuitively obvious, and sometimes an example is presented that illustrates the absurdity of violating the condition. It is hinted, but never spelled out, that to violate the condition would be a logical contradiction. I grant that in a set of particular circumstances, it may well be that a violation of one of the conditions has absurd consequences. In another set of particular circumstances, however, it may be acceptable, or perhaps even reason would demand, that choice violate one of the conditions. A condition may be substantively applicable in particular circumstances, or it may be useful as a simplification to assume that one or the other of the conditions applies, or to assume that it usually applies unless there are special circumstances. But neither of the conditions is necessary to practical reason in the sense that it should apply to each and every possible choice regardless of the particular circumstances. My goal is to show that the conditions are not requirements of rationality, are not justified by naked appeal to intuition, and to do so I present examples that illustrate the absurdity of
obeying the condition. If I present plausible counterexamples to the conditions, then
my argumentative goal is achieved.

It should be understood that all Arrow (1952, 49) intends by the word *rational*
is that “an individual is rational if his preferences among candidates can be expressed
by an ordering; similarly, collective decisions are made rationally if they are
determined by an ordering acceptable to the entire society.” Although the IIA(RM) is
implicated in the ordering assumptions of Arrow’s theorem, the independence aspect
of the IIA(A) is an additional requirement. To violate these narrow construals of
rationality is not to violate the broader concept of rationality: of having beliefs and
desires, and carrying out plans and actions, for good reasons. Acting for good reasons
is prior to Arrow’s rational-choice model, and in case of conflict it is the model that
must go.

Here is an example showing that violation of IIA(RM), contraction
consistency, may be substantively rational. My example is inspired by Sen (1993),
who argues that the concept of internal consistency of choice, exemplified by the
IIA(RM), is “essentially confused, and there is no way of determining whether a
choice function is consistent or not *without* referring to something external to choice
behavior (such as objectives, values, or norms).” To introduce and motivate his entire
scheme, Arrow (1963/1951, 2) supposes a society that must choose among
disarmament, cold war, or hot war. It is obvious to Arrow that rational behavior on
the part of the community would mean, in analogy to the economist’s understanding
of individual choice, that “the community orders the three alternatives according to its
collective preferences once for all, and then chooses in any given case that alternative
among those actually available which stands highest on its list” (1963/1951, 2). He
then uses the Condorcet paradox of voting to illustrate the possibility that the
community might cycle among the three momentous alternatives. Arrow’s scheme requires that choice among more than two alternatives be decomposed into pairwise comparisons. The example Arrow chooses, however, illustrates the folly of insisting on pairwise comparisons over social states.

Suppose that there is an individual who prefers peace so long as it does not require surrender to the enemy. If she were to face all three of Arrow’s alternatives, then she would rank them Cold War > Hot War > Disarmament. She least prefers Disarmament as that would amount to surrender to the enemy, but also thinks Cold War is better than Hot War because there are fewer casualties in Cold War. If she were to face a choice between Cold War and Hot War, she would choose Cold War. If she were to face a choice between Hot War and Disarmament, she would choose Hot War. If she were to face a choice between Cold War and Disarmament, however, she would choose Disarmament. Why? If Hot War were off the menu of choice, if Hot War were no longer possible, then the peace of Disarmament would be preferable to the tension of Cold War and would not require surrender to the enemy. Her preferences over a menu of all three alternatives is transitive: Cold War > Hot War > Disarmament. Her preferences over menus of two alternatives differ, however, and to chain them yields a cycle: she would prefer Cold War to Hot War to Disarmament to Cold War. One may not agree with the order of her rankings, but one would have to agree that her rankings are substantively rational.

Arrow’s argument about social choice is by analogy to individual choice. The Arrovian scheme seems to assume that for an individual’s choice to depend upon the menu of choices is irrational, but I have just shown by example that it is possible for a rational person’s preferences over alternatives to vary by the menu of alternatives available. If it is possibly rational for an individual’s choices to vary by menu, then,
by analogy, it is possibly rational for a society’s choices to vary by menu. A collective might rationally rank $A > B$ when those two are the only alternatives of interest, but rank $B > A$ when alternative $C$ is also available. Thus to demonstrate formally that a voting rule ranks $A > B$ when $A$ and $B$ are under consideration, but $B > A$ when $A$, $B$, and $C$ are under consideration is not in itself an objection to the voting rule. One would have to go beyond formal rationality and further show that the reversal is substantively irrational in the concrete instance. Suppose that some reversals are substantively rational and some substantively irrational. Then, in the comparative evaluation of voting rules, we would want to know the probable frequency of substantively irrational reversals for each rule as conditions vary. Substantively rational reversals would be welcome.

Now for the more important chore, to show that violation of Arrow’s IIA(A) may be substantively rational. Suppose that there is to be a reception and that the caterer will only provide one beverage, either beer or coffee. The overly rushed organizer of the reception copies a form from last year’s event that asks people to rank beer, coffee, water, tea, milk, and pop and emails it out. Attendance is by RSVP only. Five people from the business school will come, and each of them ranks beer > coffee > water > tea > milk > pop. Four people from the law school will come, and each of them ranks coffee > beer > water > tea > milk > pop. The organizer, a political scientist indoctrinated in the Arrow theorem, and a believer in the IIA(A) condition, tallies only preferences over beer and coffee, the two relevant alternatives. Five want beer rather than coffee and four want coffee rather than beer: beer wins by majority rule. Beer is the Condorcet winner, the alternative that beats all others in pairwise comparison: beer > coffee water > tea > milk > pop. Beer is also the Borda winner. It turns out though that the four people from the law school cancel, and four
people from the theology school will attend instead. Their ranking is: coffee > water > tea > milk > pop > beer. The ranking of the lawyers and the theologians is almost the same, except that the lawyers rank beer second and the theologians rank beer last.

The organizer looks only at the relevant alternatives, coffee and beer: by pairwise comparison nothing has changed, beer is still the choice by majority rule. The Condorcet order remains identical as well. The theologians come to the reception and are furious. They are teetotalers and would rather have anything but beer. The organizer loses his job, all because of his dogmatic belief in the IIA(A) condition. If the political scientist had instead used the Borda count, he would have noticed that the theologians ranked alcohol last, and would have provided the Borda-winning beverage, coffee.

### Table 3. Substantively Rational to Violate IIA(A)

<table>
<thead>
<tr>
<th>Business School</th>
<th>Law School</th>
<th>Theology School</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Beer</em></td>
<td>Coffee</td>
<td>Coffee</td>
</tr>
<tr>
<td>Coffee</td>
<td><em>Beer</em></td>
<td>Water</td>
</tr>
<tr>
<td>Water</td>
<td>Water</td>
<td>Tea</td>
</tr>
<tr>
<td>Tea</td>
<td>Tea</td>
<td>Milk</td>
</tr>
<tr>
<td>Milk</td>
<td>Milk</td>
<td>Pop</td>
</tr>
<tr>
<td>Pop</td>
<td>Pop</td>
<td><em>Beer</em></td>
</tr>
</tbody>
</table>

Business School + Law School
Condorcet: *Beer* > Coffee > Water > Tea > Milk > Pop
Borda: *Beer* > Coffee > Water > Tea > Milk > Pop

Business School + Theology School
Condorcet: *Beer* > Coffee > Water > Tea > Milk > Pop
Borda: *Coffee* > Water > Beer > Tea > Milk > Pop

The viewpoint of the Arrow theorem is that a social choice is derived from a social ordering that is aggregated from individual orderings over social states; “only orderings can be observed, and therefore no measurement of utility independent of
these orderings has any significance” (Arrow 1967, 76-77), and the IIA(A) condition enforces the ban on information other than individual orderings. What if there had been a discussion about the beverage to be served at the reception? Assume that the theologians have the same mere ordering of preferences. Before the decision, they explain that they are teetotalers, and would rather have anything but beer. Alternatively, suppose that they are prohibited from drinking beer as a matter of their religion, and that they have a right to be served a beverage at the reception that doesn’t offend their beliefs. The political scientist, our hapless believer in the IIA(A) condition, would have to reject this information and enact the social choice of beer. “Only orderings can be observed,” he would reply to the theologians, “and the only ordering that matters is between coffee and beer. That you would rather have any beverage but beer is irrelevant.”

I have put the argument in terms of the Borda count because my focus is on voting rules aggregating from ordinal preferences. In principle, a utilitarian voting rule might be applied to cardinal preferences to obtain a more accurate social ranking (maybe beer would be last), but the qualitative features of the story would remain the same. Riker (Riker and Ordeshook 1973, 110-111) has an objection to the kind of story I have told. Adapting his objection to fit the current example, what if the orderings remained the same, but the participants had cardinal utilities as follows. Each member of the business school would rank beer as 10, then the remainder of alternatives as in the Borda count, 4, 3, 2, 1, 0 and each member of the theology school as in the Borda count, 5, 4, 3, 2, 1, 0. With these cardinal utilities, although the Borda count would still select coffee, the utilitarian choice would be beer (51) over coffee (40). My first response is that in the pure voting exercise all we have are ordinal data, and thus we cannot go beyond the Borda count. But if we had sound
cardinal data then why not use it? My second response is that the Rikerian objection has changed the qualitative features of the story, now, although the theologians rank beer last, the business schoolers are crazy for beer. My third response is that if the Borda count generally is imperfect in approximating cardinal utilities, should that be a goal, then Condorcet is more imperfect at the task – in my original example Condorcet fails to detect that the theologians rank beer last. In the Rikerian example, Condorcet picks beer, but, in utilitarian terms, for the wrong reasons. My fourth response is that with increasing numbers of voters, Borda tends to the utilitarian outcome (as does Condorcet). To conclude, in practice, ardor or horror at beer is expressed in discussion, people vote in a fair-minded way take account of intense preferences of others (and maybe entrench some relating to life and liberty as rights claims), and in more competitive environments may engage in logrolling (OK, no beer, you theologians, but we get to choose the main dish).

Arrow does not defend his assumptions on logical grounds. As for the requirement of pairwise comparison for individual choice orderings,

There seems to be no logical necessity for this viewpoint; we could just as well build up our economic theory on other assumptions as to the structure of choice functions if the facts seemed to call for it.\(^{18}\)

\(^{18}\) Like Lange, the present author regards economics as an attempt to discover uniformities in a certain part of reality and not as the drawing of logical consequences from a certain set of assumptions regardless of their relevance to actuality. Simplified theory-building is an absolute
necessity for empirical analysis; but it is a means, not an end. (Arrow 1963/1953, 21)

Arrow defends pairwise comparison as only a modeling convenience in the case of individual choice. And, his justification for pairwise comparison in the case of social choice is only by analogy to individual choice. I have argued that it can be substantively rational to violate either IIA(RM) or IIA(A). There is no need to make fetishes of the conditions, they are model-building means, not normative ends. We should not let the simplifications of models mislead our larger judgements about, say, democracy.

My analysis complements one of the few critical examinations to be found anywhere of the normative authority of the IIA(A) condition:

The independence condition is certainly normative in content – it sets sharp limits on the information deemed to be normatively irrelevant and on the allowable form of that information – but what is its normative justification? In fact . . . most attempts at justifying the independence requirement are methodological rather than normative – concerned with analytical convenience and the wish to avoid the issues involved in the debate on the interpersonal comparability of welfare. But if, in practice, individual evaluations of alternatives are not always independent of further options, such methodological points have no relevance to the normative justification of the independence requirement.
Brennan and Hamlin (2000, 107-108) conclude that the IIA(A) is a strong and essentially arbitrary normative requirement, in that it does not derive from consideration of the preferences of the individuals who constitute society.

The Irrelevance Justification Is Flawed. Arrow (1952) says that the IIA(A) condition has always been implicitly assumed in voting systems, but the claim is mistaken. He offers the example of a community deciding between construction of a Stadium and of a Museum. The community can afford one or the other, but not both, and the community cannot afford at all a University. Arrow believes that the choice between the Museum and the Stadium must be independent of preferences of community members between the feasible Museum and the infeasible University. “The essential argument in favor of this principle is its direct appeal to intuition” (51). It is true, as a matter of practice rather than of logic, that infeasible or irrelevant alternatives are usually not placed on ballots. But sometimes they are. It was notorious in the 2000 election that dead Democratic candidate Carnahan remained on the ballot for U.S. Senator from Missouri, and won the election against Republican Ashcroft. As expected, the Democratic governor appointed Carnahan’s wife as his successor. In Great Britain, the Official Monster Raving Loony Party is an irrelevant alternative that regularly appears on the ballot.

In his narrative justifications Arrow uses an ordinary conception of irrelevance: voting systems should choose over available alternatives but not over conceivable alternatives, over feasible alternatives but not over possible alternatives, over relevant alternatives but not over irrelevant alternatives. I imagine that many people are against the mischief of irrelevance, and also against permitting preferences over irrelevant alternatives to influence wrongly decisions concerning relevant
alternatives. The ordinary irrelevance that Arrovians deplore in narrative
justifications is not the irrelevance formally stated in the IIA(A) condition, however.

What if the community could afford a Museum, a Stadium, or a University, but not any two or all three of these alternatives, and further could not afford at all a Nuclear Missile? There would have to be a social choice among the three feasible alternatives. Does the IIA(A) work to permit consideration of the relevant Museum, Stadium or College and forbid consideration of the irrelevant Nuclear Missile? Not at all. Arrow’s condition does not partition alternatives into the ordinarily relevant and the ordinarily irrelevant. The condition applies to all candidates \( x \) and \( y \), let’s say in a set \( S \). If there are four relevant alternatives, \( a, b, c, \) and \( d \) in \( S \), then the choice among \( a, b, \) and \( c \) must be independent from preferences involving \( d \). As we consider the three alternatives \( a, b, \) and \( c \), the choice between \( a \) and \( b \) must also be independent from preferences involving \( c \) or \( d \). The choice between \( a \) and \( c \) must be independent from preferences involving \( b \) or \( d \), the choice between \( a \) and \( d \) must be independent from preferences involving \( b \) or \( c \), the choice between \( b \) and \( c \) must be independent from preferences involving \( a \) or \( d \), and the choice between \( b \) and \( d \) must be independent from preferences involving \( a \) or \( c \), even though each of \( a, b, c, \) and \( d \) is ordinarily relevant. The IIA(A) condition always boils down to one that requires that the social choice between any two alternatives \( x \) and \( y \) not be influenced by individuals’ preferences over any third alternative. The IIA(A) would better be named the \textit{pairwise comparison condition}, as it requires that choices among several alternatives be carried out only with information about choices between pairs. Arrow (1963/1951, 20) said as much in 1951:

One of the consequences of the assumption of rationality is that the choice to be made from any set of alternatives can be determined by
the choices made between pairs of alternatives. Suppose, however, that the situation is such that the chooser is never confronted with choices between pairs of alternatives; instead, the environment may always involve many alternatives. . . . we can say that the choices made from actual environments can be explained as though they were derived from choices between pairs of alternatives; and, at least conceptually, it makes sense to imagine the choices actually being made from pairs of alternatives.

The IIA(A) means that if someone ranks \( x > y > z \), we count that she likes \( x > y \), count that she likes \( y > z \), and count that she likes \( x > z \), but we are not allowed to count that she likes \( x > y > z \). Saari (2001b) argues that the information lost due to this prohibition is what drives the impossibility result. We can insist that voting procedures rely only on pairwise comparisons and end up with Arrow’s dictatorship result and with startling interpretations such as those in my hall of quotations, or we can more sedately interpret the Arrow theorem to mean that procedures for three or more alternatives require more information than pairwise comparisons (Saari 1995a, 88).6

Continue to suppose that any one of the Museum, Stadium, or the University is a feasible or ordinarily relevant alternative, but not any two or all three. The choice between the Museum and the Stadium cannot be affected by people’s preferences between the Museum and the University or between the Stadium and the University, according to the IIA(A), even though the University is a relevant alternative in the ordinary sense of the term. Voter preferences are distributed as in Table 4.

Table 4. The Relevance of Irrelevant Alternatives

| Actual  | 99,000 voters | 100,000 voters |
Voters are asked to rank all alternatives. Begin with the “actual” scenario in Table 4. The Condorcet advocate insists that the Stadium should win, even though almost half the voters rank it last among all projects. The Borda advocate insists that the Museum should win, it is the first choice of almost half the voters and the second choice of the other half. The presence of the third alternative of the University on the ballot assists in the decision because it discloses that the 99,000 voters rank the Stadium last. In choosing between the Museum and the Stadium, eliminating from consideration the genuinely relevant alternative of the University ensures that the Stadium wins, and conceals the fact that almost half the voters would rather build anything but the Stadium. The Condorcet rule does not violate the IIA(A), but in this example the Borda rule does violate the IIA(A). Compare the pair of the Museum and the Stadium, and inspect the “counterfactual” scenario in Table 4. Suppose the 99,000 voters change their ranking of the university from second to third. Then the social choice by the Condorcet rule would continue to be the Stadium, but the social
choice by the Borda rule would change from the Museum to the Stadium – the IIA(A) is violated. Next, suppose that the university is infeasible, is an ordinarily irrelevant alternative. Nothing in the foregoing analysis changes, except that addition of the ordinarily irrelevant alternative would have made more information available for a better decision. The IIA(A) decrees that in all circumstances there is nothing to be said in favor of any method that considers information beyond pairwise comparisons. If the IIA(A) were strictly and literally applied, it would forbid the social choice process even from considering any public arguments concerning the alternatives, as that would be information beyond the pairwise rankings of voters.

The Arrovian tradition equivocates on “relevance.” The IIA(A) condition does nothing more than require that in a choice between two alternatives a third alternative should have no influence. Whether any of those alternatives are relevant or irrelevant, feasible or infeasible, available or unavailable, in the ordinary sense of those terms, has nothing to do with the IIA(A) condition. The choice could be among two ordinarily irrelevant alternatives, and the IIA(A) would forbid that a third ordinarily relevant alternative influence the choice between the two ordinarily irrelevant alternatives (then we would have to rename it the independence of relevant alternatives condition). The choice could be between ordinarily relevant alternative $x$ and ordinarily irrelevant alternative $y$, and then the IIA(A) condition would require that the social choice between $x$ and $y$ not be influenced by preferences over some third alternative $z$, no matter whether $z$ is ordinarily relevant or irrelevant.

Arrow (1952) says that all actual voting methods respect IIA(A). It is true that most elections do not consider ordinarily irrelevant alternatives (and when they do voters mostly ignore them), but it is definitely not true that all voting methods proceed by pairwise comparison. For example, suppose there is a natural election among
candidates, say there are six. There are certain qualifications for entry, such as residence and age, and to be eligible a candidate must declare before a certain date. The election is carried out by Hare preferential voting. This election does not violate ordinary irrelevance because no ordinarily irrelevant candidates are considered. It does violate IIA(A), however, because the Hare method does not proceed by pairwise comparison.

Arrow (1969) contains a remarkable statement:

For example, a city is taking a poll of individual preferences on alternative methods of transportation (rapid transit, automobile, bus, etc.). Someone suggests that in evaluating these preferences they also ought to ask individual preferences for instantaneous transportation by dissolving the individual into molecules in a ray gun and reforming him elsewhere in the city as desired. There is no pretence that this method is in any way an available alternative. The assumption of Independence of Irrelevant Alternatives is that such preferences have no bearing on the choice to be made. It is of course obvious that ordinary political decision-making methods satisfy this condition. When choosing among candidates for an elected office, all that is asked are the preferences among the actual candidates, not also preferences among other individuals who are not candidates and who are not available for office.

If the IIA(A) states that nonexistent alternatives should not be listed on ballots, then there would be no controversy about it. The IIA(A), of course, states something else entirely, that only pairwise comparisons should be inputs to social choice. Yes, the IIA(A) agrees with common sense by excluding the ray gun, but at the cost of
excluding all but pairwise voting in consideration among the feasible alternatives of rapid transit, bus, automobile, etc. The IIA(A) way overshoots. It is as though someone in Canberra refuses to leave his room because he’s heard that there’s a dangerous snake somewhere in Sydney. We point out to him that he won’t get bit by walking around Canberra, but he replies that he wishes to get no closer to the snake.

We must be careful here about confusing the IIA(A) and the IIA(RM). I don’t think that Arrow in the example is thinking of adding the ray gun to the menu of alternatives (possible violation of IIA(RM)). What he means, I think, is that preferences over the infeasible ray gun shouldn’t influence preferences over (any two) feasible alternatives (possible violation of IIA(A)). The way to avoid that influence is to decompose all social choice into pairwise comparisons, and then to string together the pairwise choices over alternatives of interest, but that remedy carries an immense price: dictatorship as the only acceptable social welfare function. It is as though the obsessive Canberran chooses to starve himself to death rather than leave his house.

In 1963, Arrow (1963/1951, 110) commented that the austerity imposed by the IIA(A) “is perhaps stricter than necessary; in many situations, we do have information on preferences for nonfeasible alternatives. It can be argued that, when available, this information should be used in social choice . . .” My business school and theology school example shows this. Later, Arrow (1997, 5) said of an approach such as Borda’s that it is not willing to take the logical next step, of adding irrelevant alternatives to the list of candidates just to get extra information. I agree that people usually would not advocate adding noncandidates in order to obtain more information, but this is a practical consideration, not a logical one. I can conceive of circumstances where people would advocate consideration of irrelevant alternatives just to gain extra information. Suppose that a forestry workers’ cooperative is voting
to select a site for a new office. The office subcommittee has searched diligently and presents to the membership the only two alternatives available on the market. One is small but centrally located, the other is remote but large. Discussion suggests that sentiment is stronger for the small office, but a straw vote over small and large indicates a tie between the two. Discussion also reveals that a large majority would prefer an intermediate alternative if it were feasible. The situation reflects the following preference orders: 25 members rank small > intermediate > large; 50 members rank intermediate > large > small; and 25 members rank intermediate > small > large. The Condorcet rule and the local Borda rule over the feasible pair yields 50 votes for small and 50 votes for large. Someone suggests voting by the global Borda count over the feasible small and large alternatives and the infeasible intermediate alternative. The result is intermediate (175) > small (75) > large (50). If all along we had used the global Borda count over the three alternatives, then we would have violated the IIA(A) with respect to the two feasible alternatives. Global Borda says that small > large, but if the preferences of the 25 who ranked small > intermediate > large were to change to small > large > intermediate then the global Borda result would change to small ~ large, in violation of the IIA(A). Adding consideration of the infeasible alternative of the intermediate office shows both that an intermediate office would be most favored, and that a small office is favored over a large office. As a result, the members instruct the office subcommittee to pursue more aggressively intermediate alternatives, and if none is found, to secure the small office. The example shows that consideration of ordinarily irrelevant or of third alternatives can be substantively rational, and is even strongly advisable in some circumstances.
Generally, though, irrelevant alternatives are not added to gain extra information. Why might that be? One reason, I suppose, is that in all contexts relevance is a compelling imperative. A stronger reason, perhaps, is that in many cases, unlike in my example of the cooperative, there is no motivation to include alternatives incapable of selection. What happens if the noncandidate wins? If the election is carried out by plurality rule, as many are, then the victory of a noncandidate would require a new election. I recall that once those of a rebellious bent proposed to place “none of the above” on American plurality ballots, but the proposal fell flat because of the practical need to have a winner. Adding irrelevant alternatives has the air of frivolity and irresponsibility. Perhaps some would be motivated to add the noncandidate in order to manipulate the election and would not vote sincerely, but certainly even if many were to consider John Stuart Mill the best alternative for mayor few would bother to vote sincerely for that choice on the ballot. Arrow’s point is one of practical observation rather than of logical objection.

**Voting Rules May Be Justified Independently of Interpersonal Comparisons of Utility.** Arrow (1952, 52) states that to violate the IIA(A) presents an “operational problem: preferences between impossible alternatives make virtually no sense, for they correspond to no action that an individual could imagine having to perform.” He must be speaking rhetorically, as immediately before he is speaking of logically possible alternatives or of imaginable alternatives, not of impossible alternatives. He must mean that individuals are required to have preferences over all possible alternatives \(X\), in order to decide among feasible alternatives in some subset \(S\), and then his argument would be that critics of the IIA(A) require people to have preferences over infeasible alternatives in \(X - S\). If he is correct that it makes no sense to have preferences over infeasible alternatives in the context of applying a voting
rule, however, then it would follow as well that Arrow’s entire social welfare perspective itself makes no sense, because in his larger scheme, “it is assumed that each individual in the community has a definite ordering of all conceivable social states” (Arrow 1963/1951, 17), which includes of course the infeasible, however that is defined. In his 1963 addendum, Arrow (1963/1951) offers further justification of the IIA(A). He begins with reference to ordinalism in economics and expresses the belief that ordinalism is desirable because it is based only on interpersonally observable behavior. The IIA(A), he says, extends the requirement of interpersonal observability a step further. One could observe all preferences over alternatives available to or feasible for society, but one could not observe preferences over alternatives not available to or feasible for society, demonstrating the practicality of requiring voting procedures to respect the IIA(A). Again, this calls into question the entire Arrovian scheme. Arrow (1967, 61) later directly mentions this inconsistency with respect to the social welfare perspective: how can we “possibly know about hypothetical choices if they are not made”? His response is to “pass by” the issue.

Arrow’s concern about observability arises from the doctrine of logical positivism, just past its apex of influence as Arrow wrote in 1951. Since 1951, logical positivism has fallen into disfavor in all disciplines outside economics, but inside that discipline its strictures linger to this day. “According to this doctrine, any statement is meaningful if and only if it is possible to specify a set of observations that would verify it; and the meaning of the statement is exhausted by the specification of these observations,” write Barry and Hardin (1982, 249).7

The viewpoint of Arrow’s Social Choice and Individual Values (1963/1951, 9-11, 109-111) is the claim that the “interpersonal comparison of utilities has no meaning and, in fact, that there is no meaning relevant to welfare comparisons in the
measurability of individual utility” (9). The only meaning that the concepts of utility can be said to have is their indications of actual behavior, says Arrow, and if such a course of behavior can be explained by a given utility function it can also be explained by a second that is a strictly increasing function of the first. Von Neumann-Morgenstern utilities over alternative probability distributions are no way out, because the behavior explained by a given function of that type can also be explained by a second that is a positive affine transformation of the first. Even if such utility were measurable, says Arrow, which function out of an infinite family should be selected to represent the individual, and which function should be selected to aggregate the individual utilities? The selection of the functions “requires a definite value judgment not derivable from individual sensations” (11). Sen (1970, 97) points out that the choice of one cardinal utility function or another is descriptively arbitrary, “but in an ethical argument one may wish to choose some particular scaling in spite of this ‘arbitrariness,’ on some other grounds that may be additionally specified.” Arrow seems to go in the other direction: “If there is no empirical way of comparing two states . . . there can be no ethical way of distinguishing them” (112). If the welfare of different individuals is empirically and hence ethically indistinguishable, then there is no reason to be more concerned for the poor than for the rich in social policy, as the rich are no better off than the poor according to the doctrine of noncomparability. Perhaps Arrow’s statement is not intended to carry outside the context of his criticism of Samuelson’s welfare function: even if the rich and poor were to be indistinguishable in terms of observed mental states, they are surely empirically and ethically distinguishable by objective measures.

In 1963 Arrow criticizes Borda’s justification of his method. Borda, says Arrow, gave equal weight to differences between adjacent rankings and gave equal
weight to different voters. “The first raises the problem of the measurability of utility, the second that of interpersonal comparability” of utilities (1963/1951, 94). Borda justified the first claim with the argument that if a voter ranks B between A and C, then we have no information about whether the intensity between A and B is more or less than the intensity between B and C, according to Arrow. The second claim “is justified on the grounds of equality of voters,” an ethical claim. Arrow immediately contrasts the Borda method to the Condorcet method (“that a candidate who receives a majority against each other candidate should be elected”) and praises Condorcet as consistent with the IIA(A).

We begin with Borda’s second claim. Notice that every democratic voting rule, including Condorcet, gives weight (usually equal) to each voter. If giving equal weight to voters in the Borda count offends the noncomparabilist doctrine, then so does the Condorcet method, and indeed any political aggregation which assigns weights, equal or unequal, to voters. In 1951 Arrow (1963/1951, 46) proved that the method of majority decision over exactly two alternatives does satisfy the conditions of his theorem, in other words, is a possible social welfare function; the impossibility result pertains to three or more alternatives (when applied pairwise over three or more alternatives we call this the Condorcet method, which can cycle). He eludes the interpersonal comparability of utility problem by means of the following definition:

By the method of majority decision is meant the social welfare function in which \( x \succ y \) [the social ranking] holds if and only if the number of individuals such that \( x \succ_i y \) [individual rankings] is at least as great as the number of individuals such that \( y \succ_i x \).

Either counting the numbers of individuals for, against, or indifferent to an alternative violates the prohibition on the interpersonal comparison of utility, and all democratic
voting is thereby rendered meaningless, or the assumption of one vote per individual has nothing to do with interpersonal comparison of utility. Reading Arrow charitably, it must be that one vote per individual has nothing to do with comparison of mental states. If one vote per individual is acceptable because it has nothing to do with comparison of mental states with respect to the method of majority decision, then it must be acceptable as well with respect to any other voting method, including Borda. Therefore, the Borda method does not necessarily rely on interpersonal comparison of utility. One variety of noncomparabilist might further urge that because of noncomparability, each voter should be given equal weight.

One argument is Borda’s first claim, as related by Arrow: just because the individual preference data are ordinal, it is best to assign equal weight to the differences between adjacent rankings. Neither the Borda count nor the Condorcet method is a utilitarian method of voting: each is an operation performed on ordinal rankings.10 That the Condorcet method (in the absence of cycles) ranks all alternatives does not necessarily make it a cardinal voting scheme, and the same goes for the Borda count. Some friends, and enemies, of the Borda count argue that it approximates a cardinal voting scheme, that it attempts to squeeze cardinal blood from the ordinal turnip (Mackay 1980, 73), but same could be said for Condorcet method. Another argument has it that as the number of voters increases, cardinal variations would tend to cancel each other out, and the Borda outcome would tend to the cardinal outcome (as would Condorcet, see Tangian 2000).

The Borda count can be independently justified on purely ordinal grounds, however. It collects exactly the same information as does the Condorcet method, as can be seen in the pairwise comparison matrix, but instead sums up the number of votes each alternative gets over every other alternative, and socially ranks by those
sums. The alternative that gets the most votes over every other alternative is the winner. There is no reference to cardinality or to comparable mental states in this justification of the Borda count. Arrow’s definition of the method of majority decision counts the number of individuals who prefer one or another of a pair of alternatives, and is phrased such that it does not necessarily have anything to do with comparable mental states. Counting the number of times an alternative is preferred by individuals to every other alternative, as does the Borda count, also does not necessarily have anything to do with comparable mental states. Objection: but that would make the outcome depend on the range of alternatives. Reply: just as the outcome should depend on the range of voters.

Any democratic voting rule – Condorcet, Borda, other – weights voters (usually equally). Once that is conceded, the Borda count, just as the Condorcet method, is otherwise justified within the noncomparabilist framework. The definition of the social welfare function requires individual and social orderings. The independence aspect of the IIA(A) is adventitious. Its purpose is to exclude voting rules otherwise compatible with ordinal input, and thereby deliver the impossibility result.

The IIA(A) Does Too Much. Riker (1982, 101) describes the IIA(A) as requiring “that a method of decision give the same result every time from the same profile of ordinal preferences.” He says that the IIA(A) “seems a fundamental requirement of consistency and fairness to prevent the rigging of elections and the unequal treatment of voters,” and that it has been disputed. In a footnote (271), he explains that the main reason for the dispute is that Arrow’s original discussion of the condition confused it with contraction consistency. The IIA(A) condition has three desirable consequences, says Riker. First, it prohibits utilitarian methods of voting.
Second, it prohibits arbitrariness in vote-counting, such as lotteries or other methods that give different social choice from the same given profile of individual preferences. Third, it prohibits, when choosing among alternatives in a set \( S \), influences from alternatives in \( X - S \). I respond in turn.

First then, we shall consider utilitarian methods of voting. Riker (1982, 118) thinks utilitarian voting “gives advantages to persons with finer perception and broader horizons.” Would utilitarian voting, which on ethical grounds counted each voter equally, and which accepted intensities of preference from cardinal rankings, were it practical to carry out and resistant to misrepresentation of intensity, be undesirable? I think the main objection to utilitarian voting is its impracticality rather than its comparable cardinality; if it were a practical voting system, it might be an excellent one (it might be an excellent voting rule, but it would not necessarily indicate what is true or right). Another advantage of the IIA(A), according to Riker, is that it forbids “the arbitrariness of the Borda count.” As I have shown, the justification of the Borda count need not depend on comparison of mental states. Furthermore, the Borda count avoids two problems of utilitarian voting: first each voter is counted equally, and second the weighting of the alternatives is established by the voting scheme and thus avoids voter misrepresentation of intensity.

Under the Borda count and with more than two alternatives, if an alternative is lowered in an individual ranking that may lower the social ranking of the alternative. Is this arbitrary? Riker (1982, 108) provides an example of how the Borda count violates IIA(A), but the example seems rather feeble, and arguably supports the opposite view that \textit{violation of the IIA(A) by the Borda count is a desirable attribute}. Suppose that Larry ranks \( A > B > C \), that Moe ranks \( C > A > B \), and that Curly also ranks \( C > A > B \). Then \( A \) and \( C \) tie with a Borda score of four points each, and are
both preferred over B which has a Borda score of one point (the Condorcet outcome from the profile is C > A > B). The IIA(A) says that the social ranking of A and C should depend only on the individual rankings of A and C. Thus, if Moe changes his ranking from C > A > B to C > B > A, then, the IIA(A) decrees, the social ranking of A should not change. The Borda count after Moe’s change, however, is four points for C, three points for A, and two points for B (the Borda outcome is now C > A > B, and the Condorcet outcome remains C > A > B). The Condorcet method reports the same outcome for both profiles of individual preferences: C > A > B. Preferences with respect to only C and A are the same in each profile, so the Condorcet method here is respecting the IIA(A) (there is no cycle in either profile). The Borda method distinguishes the two profiles. Under the first profile, the Borda method finds that C is tied with A. Why does this differ from the Condorcet outcome? Condorcet reports that C is better than A, but the Borda count takes into account that although C is first-ranked by two voters, it is last-ranked by one voter, and that A is first-ranked by one voter and second-ranked by two voters. Under the second profile, the Borda method finds that C is better than A. Why the change in the Borda outcome from the first profile to the second profile? In the first profile, C is tied with A, but the change from the first to the second profile is that one voter changes A from second-ranked to last-ranked. The Borda method responds to this change, but the Condorcet method does not so respond. The IIA(A) says that a voting rule should never respond to that sort of change. Practical evaluation of alternative voting rules must consider actual frequencies of good and bad violations of the IIA(A).

Second, the IIA(A) also excludes any probabilistic voting rule, according to Riker (1982, 118), because from the same given profile of individual preferences it requires that a voting rule always return the same result. For example, imagine the
rule that each of \( n \) voters writes down his first preference among the alternatives on a ballot and then one ballot is randomly drawn to determine the social choice. Incidentally, such a rule would be strategy-proof, since each voter has the incentive to report her true preference, which is why the Gibbard and Satterthwaite theorems are limited to definite voting rules. Or suppose a rule such that if 90% of voters prefer \( x \) over \( y \), there is a 90% chance that the social choice would be \( x \) over \( y \); such a rule would be prohibited by the IIA(A). The missing premise in his argument is that Riker does not establish that probabilistic voting rules are normatively undesirable in all or in any circumstances. We would have to consider in detail, theoretically and practically, the consequences of various such rules.

Third, “many people believe that judgments on alternatives in \( X - S \) are germane to judgments on \( S \) itself,” that is, they doubt the IIA(A) (Riker 1982, 129). To charitably reconstruct Riker’s argument, there is no formal method to decide actual relevance and irrelevance, thus, the only way to guard against the threat of irrelevant alternatives having an influence on decisions is by way of the pairwise IIA(A) condition. More precisely, there is no “wholly defensible method to decide on degrees of irrelevance. In the absence of such a method, Condition [IIA(A)] seems at least moderately defensible.”

If, for whatever reason, one were motivated to satisfy Riker’s first and second concerns, to disqualify utilitarian or probabilistic voting methods, then other approaches are available that would do the job yet would also permit some voting systems; and if there were somewhere a real human being concerned about wrongful influences from consideration of ordinarily irrelevant alternatives, then there are approaches that address that concern but avoid the impossibility conclusion.
I begin with Hansson (1973), of whom Riker (1982, 275) is aware. If feasible and infeasible alternatives can be clearly identified, then one should simply confine consideration to feasible alternatives. For example, if the application of interest is a standard political election among candidates, then the feasible candidates are those who stand for the election. Infeasible candidates would not be considered just because they are infeasible candidates. Then the election would be independent of infeasible alternatives, not in Arrow’s sense of avoiding consideration of third alternatives, but rather in the ordinary sense of avoiding actually infeasible alternatives. Problem solved. Nor is this terminological trickery: the natural reaction to any proposal to include, for example, Napoleon and Confucius as extra candidates in the city council election, is ridicule and rejection. This is a clear and simple solution to the IIA(A) problem with natural applicability in the political context. Because it is so simple, I take a moment to reiterate that people sincerely concerned about consideration of infeasible alternatives have no worries in this political context.

One of Arrow’s (1952, 52) main concerns with respect to the IIA(A) is:

If it is abandoned, a choice among a given set of candidates can be made only if each individual possesses a list of preferences containing more candidates than those which are really available. What will be the list of candidates that the individual will be asked to order? There is no natural limit except a vague universe comprising all logically possible candidates. It does not appear correct to make the choice among a very limited set of possibilities depend on all preferences among all “imaginable” possibilities.

Notice that this objection does not apply to the voting context discussed in the last paragraph. Perhaps it does apply in the social welfare context, and perhaps welfare
economists have more reason to worry about the IIA(A) than those in political studies. The worry is that the social welfare function would have to consider not only actual alternatives, but also the analogs of Thomas Jefferson, Socrates, Buffalo Bill, fictional people, animals, fictional animals, androids, and rocks, and “it is costly in terms of resources to gather and process information about preferences – independence is a requirement that we conserve those resources” (Kelly 1988, 73).

First, the costly resources argument is surprising. I think of the social welfare function as an idealization, not an exercise actually carried out in full. Decision theory constantly proposes simplifying models that it would be practically impossible for real agents to carry out, for example, because of combinatorial explosions. The practical impossibility of carrying out the calculation is seldom raised as an objection to the model. Second, it may well be, and I conjecture that it is likely, that the influence of irrelevant alternatives would be beneficial, neutral, or trivial, especially with large numbers of voters and alternatives. Third, I am not aware of whether welfare economists have explored the concept of possibility with more philosophical rigor; if not, perhaps that would help clarify the problem.

Hansson offers weakened independence conditions. Again, $X$ is the set of all alternatives, nonempty subset $S$ contains the feasible alternatives, and $X - S$ contains the infeasible alternatives. IIA(A) would be violated if changes in preferences over two alternatives in $X - S$ changed the social choice over alternatives in $S$, and also would be violated if changes in preferences over two alternatives, one in $S$ and one in $X - S$, changed the social choice over alternatives in $S$. Hansson’s “positionalist independence” (PI) condition, like IIA(A), forbids changes in preferences over two infeasible alternatives in $X - S$ to influence the social choice, but it permits changes in preferences over two alternatives, one in feasible $S$ and one in infeasible $X - S$ (or both
in $S$), to influence the social choice over the alternatives in $S$. This PI condition would void Arrow’s impossibility result for the Borda count, the Copeland method, and some other voting rules.

He also offers a “strong positionalist independence” (SPI) condition. How much can you change individual orderings without changing the social choice between two alternatives $x$ and $y$? Divide an individual’s rankings of alternatives other than over $x$ and $y$ into five groups: those below both $x$ and $y$, those the same as $y$, those between $x$ and $y$, those the same as $x$, and those above both $x$ and $y$. IIA(A) permits that we can move around in individual orderings alternatives other than $x$ and $y$ as much as we like, from any group to another, without changing the social choice. The PI condition allows us to move around alternatives other than $x$ and $y$ only within each of the five subgroups. SPI is more generous than PI, it allows us to move alternatives from above both $x$ and $y$ to below both $x$ and $y$, and from below to above, without changing the social choice. Although stronger, the SPI condition still permits the Borda count, that is, substituting SPI for IIA(A) in the Arrow theorem shows that the Borda count is a possible social welfare function (but not the Copeland rule, etc.)

Much more simply, Saari (1995, 97) offers an “intensity IIA” condition. Suppose that we measure intensity of preference only by the number of candidates a voter uses to separate a pair, as in the Borda count. The intensity IIA requires that the ranking of any two candidates depends only on each voters’ ranking of those candidates and the intensity of that ranking. Again, the intensity IIA avoids Arrow’s impossibility result. Also, recall that Young-Kemeny, a Condorcet-extension voting rule which resolves cycles, gets by with a local independence of irrelevant alternatives condition.

The alternative independence conditions would entirely satisfy Riker’s first and second concerns to forbid utilitarian and probabilistic voting rules. The device of
considering only feasible alternatives when there is a clear boundary between the feasible and the infeasible, such as in many democratic-voting applications, entirely satisfies Riker’s third concern about avoiding influence from preferences over infeasible alternatives, even though it would violate Arrow’s IIA(A) if more than two feasible alternatives were considered. When the boundary between the feasible and the infeasible is less clear, the weakened independence conditions partially satisfy Riker’s concern that decisions be independent of infeasible alternatives, a concern about which he can say no better than that it is “moderately defensible.”

**Independence Is Not A Practical Requirement.** Recall that Arrow believes that a positional method such as the Borda count is defective because it must consider some vague universe of all possibilities. The Borda method must consider all possible candidates, not merely actual candidates. The Condorcet method is independent of irrelevant alternatives, therefore it need not consider all possible candidates, but rather may consider only actual candidates, and thus is practically advantageous, the story goes. The Arrovian tradition has added that the Borda method is subject to manipulation by the addition and subtraction of candidates, but that the Condorcet method is not. Mueller (1989, 394) glosses Arrow:

> The outcomes under the Borda procedure and similar schemes depend on the specific (and full) set of issues to be decided. Thus, abandonment of the independence axiom raises the importance of the process that selects the issues to be decided in a way that its acceptance does not.

The supposed contrast is merely academic folklore, however. Notice that the violation alleged is of the IIA(RM), not the IIA(A). The IIA(A) takes alternatives considered as given, and varies preferences over alternatives; the IIA(RM) takes
preferences as given, and varies alternatives considered. Both the Condorcet and the Borda methods depend, although in a tenuous way, on the set of issues to be decided. And both the Condorcet and the Borda methods are susceptible to manipulation by addition or deletion of candidates.

The Borda count is theoretically subject to a special case of manipulation (Dummett 1998) – by adding new candidates from outside the given set a manipulator might be able to change the outcome. Let’s begin with a profile of five voters and two alternatives, and the accompanying pairwise comparison matrix, as in Table 5.

Table 5. Borda Manipulation, Initial Situation

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>Y</th>
<th>Borda</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Obviously, X is the Borda winner. If the Borda count is the voting rule, then supporters of Y can boost Y to first place, however, by introducing an alternative Z that is very similar to Y but just below it in everyone’s preference rankings. See Table 6.

Table 6. Borda Manipulation, First Step

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>Borda</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>
This little trick changes the Borda order from $X > Y$ to $Y > X > Z$, a violation, of course, of contraction consistency. Dummett proposes a fix for this eventuality that shall not detain us here. Another fix is strategic deterrence: the partisans of $X$ can restore $X$ to first place with an identical maneuver, they just introduce alternative $Q$ that is similar to $X$ but just below it in everyone’s rankings.

**Table 7. Borda Manipulation, Second Step**

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>Q</th>
<th>Borda</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

So we are back again to $X > Y$, just where we started. Why would the partisans of $Y$ introduce manipulative alternative $Z$, if they anticipated that the partisans of $X$ would respond with the introduction of countermanipulative alternative $Q$? The result if everyone manipulates is the same as the result if no one manipulates, and if there were transaction costs to adding alternatives then would anyone bother with this ploy?

Assume, however, that the Borda method is guilty as charged. The same charges can be laid against the Condorcet method. In order to carry out a proper Condorcet ranking one has to rank all alternatives feasible and infeasible. Why?

Suppose that the two feasible alternatives are $x$ and $y$, and that the social preference over the set $S$ of $x$ and $y$ is $x > y$ and thus that the social choice appears to be $x$. That is not sufficient to determine the social choice, however. We also have to examine *all*
infeasible alternatives $z$ in $X - S$ in order to determine that it is not the case that the social preferences are $z > x$ and $y > z$. Because if it happens that for some alternative $z$ that $z > x$ and $y > z$, then there is a cycle $z > x > y > z$. If there is such a cycle, then the choice set is empty. When we apply the Condorcet method only to feasible alternatives $x$ and $y$, the choice between $x$ and $y$ is $x$. When we apply the Condorcet method to all alternatives feasible and infeasible, then it may be that there is no social choice between $x$ and $y$. The Condorcet method seems to be especially absurd here: society would prefer relevant $x$ to relevant $y$ despite irrelevant alternative $z$, but the possibility of irrelevant alternative $z$ prohibits choice between relevant $x$ and relevant $y$. If the objection is tenuous against the Condorcet method, then the objection is tenuous as well against the Borda method.

How do the Condorcet and Borda methods compare with respect to the manipulative addition and deletion of alternatives? Someone who would manipulate by addition or subtraction of alternatives needs a vast amount of information. She needs to know in advance, before the vote, the preference rankings of all of the voters. If she is a manipulator by way of addition of alternatives, which I will show is the more important case, she also needs to know not only the voters’ preference rankings over the “relevant” alternatives naturally under consideration but also their rankings over some number of “irrelevant” alternatives in order to be able to select from the “irrelevant” alternatives the one or more that would permit her manipulation to succeed (assuming that the one or more irrelevant alternatives she needs exists). Again we stumble across an inconsistency in Arrovian tradition. On the one hand, change of preference over an “irrelevant” alternative may suspiciously change the outcome. On the other hand, the ordinalist revolution requires that welfare judgments be based only on interpersonally observable behavior, according to Arrow
(1963/1951, 110), “that, ideally, one could observe all preferences among the available alternatives, but there would be no way to observe preferences among alternatives not feasible [irrelevant] for society.” On the one hand, the possibility of manipulation is a threat. On the other hand, the information required for the manipulation is not available.

If an aspiring manipulator has the requisite information, then a new problem will arise in many circumstances. With more than a few voters and a few alternatives it is quite complicated to calculate the manipulative strategy. The calculation of a strategy may take such a long time that it is not practical to carry out, and in the extreme it may be exponentially hard to carry out, that is, it would take more time to calculate than there is in the universe, for example. Bartholdi, Tovey and Trick (1989) showed that for more than a few voters or alternatives it would be computationally impractical (NP-hard) to calculate the election outcome by the Dodgson method or the Young-Kemeny method. Bartholdi, Tovey and Trick (1992a) showed that a version of the Copeland method is computationally resistant to manipulation by strategic voting; Bartholdi and Orlin (1991) showed the same for the single transferable vote. Bartholdi, Tovey and Trick (1992b) compared the manipulability of plurality rule to the Condorcet method with respect to adding candidates, deleting candidates, partitioning candidates, adding voters, deleting voters, and partitioning voters. Although plurality is logically susceptible to the first three manipulations, it is computationally resistant to such manipulations; Condorcet is logically susceptible to the second three manipulations, but is computationally resistant to them. There is the magnificent and crisp issue of whether a manipulation is calculable in the lifetime of the universe that Bartholdi and coworkers consider, and there is also the modest and fuzzy issue of whether a manipulation is practically
calculable by real humans in real time in real settings, that could only be decided by empirical investigations.

It is important to distinguish whether the option of manipulation by addition and deletion of candidates is open only to one actor or is open to all. If it is open to one unconstrained monopoly actor, then no matter what reasonable voting rule is in force, she can just delete all candidates she doesn’t like and add the one candidate she does like, which is equivalent to agenda dictatorship. Practically, the addition of alternatives must be relatively open, and the subtraction of alternatives must be relatively closed (or made on the basis of criteria selected well in advance under veil-of-ignorance conditions, otherwise one would face the conundrum of needing to apply the voting rule in order to decide which candidates to subtract from consideration by the voting rule). If we posit that an actor has a manipulative advantage arising from her unconstrained monopoly power to delete candidates we have not demonstrated that a voting rule has unfair consequences. We have said nothing different than it would be unfair to give 99 of the voters one vote each while giving one special voter 100 votes to cast; we have quietly assumed violation of voter equality in order to publish loudly a violation of voter equality. Perhaps it is meant that a trusted and accountable authority would delete a candidate in order to manipulate by stealth.

Suppose that information and calculability are not problems. Define the criterion of manipulation to be changing a winner into a nonwinner by the addition or deletion of alternatives. Then the Condorcet method is just as vulnerable to manipulation by deletion of candidates as is the Borda method. The manipulating Condorcet chairperson simply deletes all alternatives that in pairwise comparison would defeat her favored alternative. It may be objected that what is meant is that the manipulator would be able to succeed by stealth under the Borda method but not
under the Condorcet method. That smuggles in a new assumption, however: not only would the manipulator have monopoly control over the deletion of alternatives, the manipulator would also have a monopoly control over information about the distribution of preferences and over calculative capacity. Then all we have shown is that those with unfair power have unfair influence. Manipulation by deletion of alternatives does not seem to be a plausible scenario.

The Condorcet method is as vulnerable to manipulation by addition of candidates as is the Borda method. If in the absence of manipulation the social choice would have been $x > y$, the Condorcet manipulator – as with Borda, any of the voters, not only the chairperson – simply adds an alternative $z$ that creates a cycle $x > y > z > x$, and $x$ is no longer the Condorcet winner. It may be objected that there may not exist an alternative $z$ that would allow for the manipulation, but the same could be said for the Borda count, that there may not exist an alternative the addition of which would change a winner to a nonwinner. We have established that both Condorcet and Borda are susceptible to manipulation by addition or deletion of alternatives.

Finally, there is an additional complexity. Manipulation begets countermanipulation. Someone contemplating the strategic addition of an alternative also has to consider the strategic response of other actors. It may be that strategic addition by all actors would cancel out and the outcome would be the same as in the absence of strategic action, it may be that strategic addition by all actors would lead to an outcome intended by no one, it may be that strategic actions are impractical to calculate. These issues are not well settled.

Gibbard (1973) and Satterthwaite (1975) showed that voting procedures that satisfy the IIA(A) are immune from strategic voting, but those that violate the IIA(A) are susceptible to strategic misrepresentation of preferences. This breathes new life
into the IIA(A) and the Arrow theorem. While admitting the attractive qualities of the Borda count, Mueller reports that “its Achilles heel is commonly felt to be its vulnerability to strategic behavior” (120). But the Gibbard-Satterthwaite theorem shows that all voting methods of interest share this defect of being susceptible to manipulation by strategic misrepresentation of preferences, including Condorcet pairwise comparison among three or more alternatives (but only if the profile of voters’ preferences is such to yield a cycle, which I have argued is rare). It may be that the Borda count is more practically and irremediably susceptible to strategic misrepresentation of preferences than other reasonable voting rules; it may be not (see Chamberlin 1985, Nitzan 1985, Saari 1990).

**Conclusion.** The Arrow theorem is a great piece of work. It illustrates an abstract limit case. It is a logical exercise, it does not describe the real world. The conditions of the theorem, especially IIA(A), are methodological assumptions, with no descriptive or normative force of their own. That a democratic voting rule violates IIA(A) is nothing to be feared, as such violation has not been established to be normatively undesirable, all the more so given that insistence on the conditions leaves the dictatorship of one as the only possible voting rule.

Any democratic voting rule forces a comparability assumption, or requires a nonwelfarist justification of voting weights, and a democratic voting rule, whether Condorcet or Borda, assigns equal influence to each citizen. If there is someone who believes that the idea of cardinal ranking of alternatives is unintelligible, or judges that it is impractical, she still might be able to endorse either the Condorcet or Borda methods, which work on ordinal rankings. If there is someone who fears the threat of Condorcet cycles, she still might be able to endorse the Borda count as a method that avoids cycles. If there is someone who worries about influence from irrelevant
alternatives, she can recommend that elections only consider actual candidates, and when the boundary between feasible and infeasible is fuzzy she can recommend the Borda count as a method that satisfies a slightly weakened independence condition. These, and other considerations in this chapter, are not a plea for the widespread adoption of the Borda count. Natural similarity among preference rankings means that in most circumstances the differences in outcomes among the reasonable voting rules are practically minimal. If there is someone concerned about manipulation by strategic misrepresentation of preferences, she might choose the single transferable vote, which experience has shown to be practically unmanipulable in this regard. Even our most inaccurate voting rule, plurality, might be desirable because of its simplicity. The alleged irrationalities of voting are greatly exaggerated.

It is time for a shift from constitutional “physics” to constitutional “engineering.” Newton’s first law of motion, another great intellectual achievement, says that an object in motion tends to stay in motion and an object at rest tends to stay at rest, unless the object is acted upon by an outside force. An inattentive interpreter of physics may advise engineers that it is futile to try to make anything go, or try to make anything stop, because of Newton’s first law. The inattentive interpreter has failed to notice the qualifying clause: unless the object is acted upon by an outside force. Arrow’s theorem says that there are no social welfare functions that satisfy his conditions. Yet we observe that there are social decision processes (discussion, institutions, voting) that work satisfactorily to translate what individuals would prefer into democratic social outcomes, and there are social welfare functions (voting) that work satisfactorily in practice. That means that they do not satisfy one or another of the Arrow conditions. That does not make democracy impossible, irrational, arbitrary, meaningless, and the rest.
There is another important issue of constitutional engineering. It is no longer enough to show that a voting rule is logically susceptible to manipulation by addition of alternatives, deletion of alternatives, addition of voters, deletion of voters, partitioning of voters, agenda control, strategic voting, and so on. It is no longer of interest to interpret such logical susceptibilities as relevant to global normative judgements about democracy, the trick is just worn out. The era of destructive social choice theory is past. Rules that are logically susceptible to manipulation may not be practically susceptible at all, some rules may be more practically susceptible than others, susceptibility may vary with the actual distribution of preferences in the population, the strong logical susceptibility of one rule might be more practically remediable than the weak logical susceptibility of another rule, and so on. The constructive social choice theory of the future would work on the comparative practical susceptibility of rules to unfair manipulations, and the identification or invention of institutions to remedy them, rather than on denunciations of democratic choice.

1 Keith Dowding, Christian List, and Bruno Verbeek saved me from many major and minor errors in an earlier draft of this chapter. They are not to blame for those that remain.

2 Arrow’s Condition I is the conjunction of two conditions that can be written separately: one requiring ordinal measurability and no interpersonal comparisons, and another purely the independence of irrelevant alternatives.
3 Arrow’s book was dated 1951, and he published a 1952 article summarizing its findings. The 1951 book was reprinted in 1963, containing an important addendum which updates and responds to critics.


5 My example is adapted from Goodman and Markowitz (1952).

6 Sen (e.g., 1982, 330) blames the exclusion of non-utility information and the exclusion of any utility information involving interpersonal comparisons for the impossibility result. Both Saari and Sen finger the exclusion of available information as the culprit.

7 Christian List (2000) argues that even within the verificationist framework, the empirical “meaninglessness” or underdetermination of interpersonal utility comparisons does not imply the impossibility of such comparisons.

8 A positive affine transformation is \( f(x) = a + bx \), where \( a \) and \( b \) are real numbers and \( b > 0 \).

9 Social choice rules such as – do whatever the Bible, or the leader, says to do – need not weight voters.

10 The ranking computation of the Borda count “is a purely formal operation on ordinal comparisons and should not be interpreted as a cardinal utility” (Kelly 1988, 71).

11 The Condorcet case, however, does not violate either of the independence conditions.