What *Should* "Majority Decision" Mean?

Michel Balinski and Rida Laraki
CNRS and École Polytechnique

Introduction

"The legitimating force of the majority rule is so pervasive that we often do not notice it and rarely do we question it: We usually take it for granted." So begins a recent article that goes on to state that majority rule "is much too powerful to make it vulnerable to a philosophical challenge" ([27], pp. 39, 40). It is our contention that majority rule *as practiced* has all too often failed to make the choice wished by the majority of voters (or members of a jury) in sometimes crucially important circumstances. Granted this fact – for which evidence is provided – there is no other course than to challenge the current philosophical view of what constitutes a majority decision and to propose a more solid alternative.

But what exactly is it? What *is* (to use an equivalent expression more congenial to this book) "majority decision"? *Majority rule* or *decision* should be, we believe, the answer to a specific, operationally pertinent question with which more than 50% of a jury or electorate can and does agree. The question posed is absolutely essential.

Every democratic country and institution pretends to elect its representatives by a majority "principle," though the precise rules by which it does so vary. The rules or methods of voting that are used differ in two ways: (1) how voters express their opinions – the *inputs* – and (2) how the various opinions are amalgamated into a collective judgment – the *output*, the numbers that measure the candidates' "standing" and determine their order-of-finish, the victor the first in the order. Every one of the methods is meant to be – and is commonly referred to being – a "majority decision."

"Majority decision" began at least as long ago as the public meetings of Greek antiquity that saw citizens raise their hands and majorities of hands decide the outcomes. In 1299 Ramon Llull [18] proposed a generalization of Condorcet's method (mainly known today as Copeland's method [13]) and in 1433 Nicolas Cusanus proposed [19] what is now best known as Borda's method. These methods both depend on the idea that voters compare pairs of candidates, answering the question: which of the two candidates do you prefer? Llull wrote: "Firstly, … two [candidates] should leave the hall. And afterwards [somebody] inquires of all others under oath which of the two is better suited and worthier…," a procedure followed for every pair candidates [18]. Cusanus wrote: "In the name of God [each elector] should ponder, directed by his conscience, who among all candidates is least qualified, and [assign him] the number One. Thereafter he should decide who is next least suitable, and [assign him] the number Two. Thus he continues until he arrives at the best, in his judgment, …" [19]. Both went on to explain how these inputs should be amalgamated, the "reasonableness" of which resided in their descriptions (much as today, in all walks of life, people propose totally ad hoc

methods that are often accepted simply because they involve counts and are, since they yield numbers, accepted as valid).

The serious scientific study of methods of voting began shortly before the French Revolution with the Chevalier de Borda in 1770 [9] and especially with the Marquis of Condorcet in 1785 [12]. Each (ignorant of the writings of Llull and Cusanus) proposed a method today associated with their names. Both were objecting to "majority decision" as then practiced, what is today referred to as *first-past-the-post*: each voter designates one candidate (in a field of two or more), the candidate most often designated wins. Borda wrote: "It is generally accepted, and to my knowledge never challenged, that in an election the greatest number of votes always designates the will of the electorate…But I will show that this opinion, that is true when the election is between only two candidates[*], can mislead in all other cases" [9] (his example is given below). Condorcet concluded, having already given his reason: "From [the example] it is easy to see that the method used in ordinary elections is defective" ([12], p. lviii).

Both called for more information from voters, namely, what Llull and Cusanus before them had asked for: an ordered list of preference over all candidates. For Borda [9], "A good method of election must give voters the means to pronounce himself on the merit of each candidate, compared to the merit of each of his opponents." Condorcet [12] asked for the same, "That each voter pronounce himself by a comparative judgment between every pair of candidates." Since then social choice theorists have studied and sought to characterize these methods and many others, in the quest of a satisfactory solution, invariably assuming that voters *compare candidates*. No really satisfactory method has emerged (as may be deduced from the continuing contending claims for one or another method). Today several methods appear to dominate in the attention they are given: Condorcet's; Borda's; Dasgupta-Maskin's [14,15] (a combination of Llull's and Borda's); and approval voting, formally introduced by Robert Weber in 1977 [31] (though it seems to have been practiced in the Sparta of antique Greece [17]).

Not one of these methods finds application in political elections. The most used methods are first- and *two-past-the-post* – the second the same as the first except that when no candidate has an absolute majority of votes, a run-off between the two front-runners designates the winner. The first is practiced in the United Kingdom and the United States, the second in France. Australia and Ireland have for years used a rather opaque method known as the alternative vote (also known as instant runoff voting, IRV), and it has recently been adopted by San Francisco for a number of citywide elections.

These methods rely on answers to different questions, so have different inputs and in practice different ballots. First-past-the-post asks for very simple expressions of voter opinion, namely, the designation of exactly one candidate (at most): which do you like best? Approval voting permits a bit more, the voter may designate as many candidates as she wishes: which

[*] An assertion that we challenge below.

would you accept? Condorcet, Borda, Llull, Dasgupta-Maskin, and the alternative vote ask the voter to submit a rank-ordering of the candidates from best to worst: what is your list of preference over all candidates? Naturally enough the procedures differ also in the way by which the votes are amalgamated or counted.

A method of voting – a realization of "majority decision" – should meet certain minimal, fundamental requirements. In our opinion they are:

- Candidates and voters must be treated equally.
- A candidate "preferred" by all voters must be the winner.
- There must be an unambiguous winner, save for rare ties.
- A winner cannot change when some other candidate drops out of the running.

The first three criteria cannot be challenged and are universally accepted. The fourth – proposed by Kenneth Arrow in 1951 [1] (in different terms that lead to the same conclusions [6]) – has been contested. It is essential, we believe, since it is so often violated in elections (with obviously important consequences) and juries.

Other "essential" requirements have been formulated. The most prominent among them is:

- A candidate preferred by a majority against every other candidate in a head-to-head race must be elected (*Condorcet-consistency*).

Such a candidate is called a *Condorcet-winner*. This principle has a major drawback: it is possible that there is no Condorcet-winner. This is the famous *Condorcet paradox* (see below for an example). Nevertheless it has occupied a central place in the psyches of voting theorists because of a belief that when there *is* a Condorcet-winner she *must* be the winner: thus the desire for a Condorcet-consistent method, one that guarantees this whenever she exists.

This belief, however, harbors serious difficulties. The fact is that neither first- nor two-past-the-post, nor Borda's method, nor approval voting, nor the alternative vote are Condorcet-consistent when voters express themselves honestly. It is also the fact that when voters are assumed to vote strategically, no method (including Condorcet's) is sure to elect the Condorcet-winner with honest votes: to elect the Condorcet-winner (assuming she exists) voters must "manipulate" or vote strategically (see Balinski-Laraki [6], chapt. 20). What manipulation means depends on the question posed the voter: when it is which one candidate he prefers, it means he may name another; when it is to name several he accepts, it means he may name some he does not accept or not name some he does accept; when it is to rank-order the candidates, it means he may submit a different order. Why? Because in so doing the voter calculates this will better serve her objectives in view of all the contingent information she has concerning the possible outcome.

To impose Condorcet-consistency decrees the Condorcet-winner *must* be elected when she exists and assumes that she is necessarily *the winner wished by the electorate*. And yet, when

she exists, there are real examples that challenge this conclusion even when there are only two candidates. For it is entirely possible that a very narrow majority of voters has a very slight preference for candidate *A* over *B* but that otherwise voters strongly prefer *B* to *A*. In this case the Condorcet-winner imposes a divisive choice instead of a consensual one (a real instance of which is given anon).

So – using (as did Lenin) the apt title of N. G. Chernyshevsky's influential 1863 novel – *What is to be done*? Still more precise information concerning voters' opinions should be sought than that suggested by Borda and Condorcet. We propose to judge candidates by evaluating them in a scale of well-understood levels of merit. Voters in a U.S. presidential election (for example) would be asked to answer a question concerning each and every candidate, "How do you evaluate her/him: *Excellent, Very Good, Good, Acceptable, Poor*, or *to Reject*?" The majority determines each candidate's merit – her/his *majority-grade* – and the majority-grades together with the relative support for them rank-orders the candidates, the winner the one in first place. One round suffices. The method – *majority judgment* – is the only one that meets what we argue are the most important requirements that a method should meet (for a complete account see [6]).

Majority decision failures

First-past-the-post measures badly. It obliges a voter to designate one candidate though he typically has much subtler appreciations over all candidates that he is unable to express. For example: some voters may consider several candidates to be equally good, others may have little regard for the candidate for whom they voted, some may vote strategically for the least objectionable candidate among those that (he calculates) have a chance, still others may simply vote for the worst candidate to express an utter rejection of all. Yet each of these votes – bearing very different meanings – is counted the same: it contributes "1" to a total. It is not in the least surprising, then, that these totals – accepted as real measures of political support by all concerned – are very inaccurate approximations, and that they often err in the "majority decisions" rendered.

The consequences of first-past-the-post's poor measures are much more serious than minor differences in relative support. Borda and Condorcet pointed out with examples that the Condorcet-winner may easily be defeated. This is all too often the case in practice. (1) George W. Bush would have lost the U.S. presidential election of 2000 if Ralph Nader had not been a candidate in Florida: there is little doubt that most of Nader's 97,000 votes would have gone to Al Gore, giving him the state's 25 electoral votes and making him the victor in the Electoral College with 291 votes to Bush's 246. (2) George Bush (his father) would have won the 1992 U.S. election against Bill Clinton if Ross Perot had not been a candidate. (3) Lionel Jospin would have reached the second round of the 2002 French presidential election – and it seems likely would have been elected in the second round – if Christine Taubira (who had 2.3% of the first-round votes) had not been a candidate. (4) François Bayrou – eliminated in the first-round – would have been elected in France's 2007 presidential election had he confronted any of the other nine contenders in a head-to-head race. (5) And in the French

2012 presidential election the rare face-to-face polls showed François Bayrou finishing neck-to-neck against the eventual winner François Hollande but well ahead of Nicolas Sarkozy, yet saw him eliminated in the first-round (winning but 9% of the votes to Sarkozy's 27% and Hollande's 29%). These real examples – and there are many others – all show how badly first-past-the-post measures. They show also how the presence or absence of candidates may change the outcome of an election: they are instances of *Arrow's paradox*.

Comparing several candidates

Borda's example against first-past-the-post assumes – as he called for – that voters have in mind rank-orders of the candidates (his example is here transformed into percentages, where 5% prefer *A* to *B* and *B* to *C,* so also *A* to *C*):

5%: *A*>*B*>*C* 34%: *A*>*C*>*B* 32%: *B*>*C*>*A* 29%: *C*>*B*>*A*

Candidate *A* wins with 39%=5%+34% of the votes to *B*'s 32% and *C*'s 29%, so the total votes give the order *A*>*B*>*C* with first-past-the-post. Yet *C* is the Condorcet-winner defeating *A* with 61% of the votes and *B* with 63%. With two-past-the-post *C* would be eliminated in the first-round and *B* would defeat *A* with 61% of the votes, giving the order *B*>*A*>*C* (this was the fate of Bayrou in the 2007 French presidential election).

But should *C* – the Condorcet-winner of this example – really be considered the choice of the majority? Different majorities support her against each of her opponents: no one majority supports her against all of her opponents, only 29% agree she is preferable to both other candidates. If the question asked of voters is which candidate do you prefer, there is no majority decision. This is the case in many first-past-the-post elections. So, in fact, the concept of a Condorcet-winner is not ideal, it is but a kind of second best notion of majority decision.

A slight change in the example shows this second best majority decision concept may not be realized either – and is an occurrence of the Condorcet paradox:

5%: *A*>*B*>*C* 34%: *A*>*C*>*B* 32%: *B*>*A*>*C* 29%: *C*>*B*>*A*

In head-to-head encounters *A* easily defeats *C* with 71% of the votes, *C* easily defeats *B* with 63%, and *B* easily defeats *A* with 61%: there is no Condorcet-winner. Actual occurrences of the Condorcet paradox have been identified in elections [21], in wine competitions [4], and in figure skating competitions [5].

To overcome the errors induced by first-past-the-post, Borda proposed a method that takes into account all of a voter's preferences. It may be described in several equivalent manners, of which the following is one. Each candidate is given a score equal to his/her average of the face-to-face votes against all of his competitors. Thus, in the second example *A*'s *Borda-score* is 55%=(39%+71%)/2, *B*'s is 49%=(61%+37%)/2, and *C*'s is 46%=(29%+63%)/2, giving the

order $A>B>C$. But Borda's method suffers from the Arrow paradox because if C were not a candidate then Borda's method elects B.

Condorcet opposed Borda's method because it is not Condorcet-consistent: when there is a Condorcet-winner it does not necessarily place her in first place. He proposed a much more complex method that also takes into account all of the voters' preferences. Instead of assigning scores to candidates he assigns scores to rank-orders of the candidates, the final outcome being the rank-order with the highest score. A voter contributes k points to the *Condorcet-score* of a rank-order if the voter agrees with it in k pair-by-pair comparisons. Thus, for example, if a voter's preference-order is $C>B>A$ she contributes 0 points to $A>B>C$, 1 point to $A>C>B$, 2 points to $C>A>B$, and 3 points to $C>B>A$. In the second example the rank-order of $A>C>B$ has a Condorcet-score of 173=10+102+32+29: 2 from a voter with preference $A>B>C$, so 10% from all with that order; 3 from a voter with preference $A>C>B$, so 102% from all; 0 from a voter with preference $B>C>A$; and 1 from a voter with preference $C>B>A$, so 29% from all. $A>C>B$ is the order with the highest Condorcet-score, so is taken as the final order of finish. Whenever there exists a Condorcet-winner he will necessarily be the winner with this method. However, the method becomes very complicated when there are more candidates, it provides no measure of the relative standings of the candidates, and it also suffers from the Arrow paradox (if C is not a candidate B is the winner, not A). Condorcet finally abandoned it to champion the one simple cause of a Condorcet-winner, even though this might mean no outcome. Henceforth this last choice is referred to as *Condorcet's method*.

The *alternative vote* – that also asks voters for a rank-order of the candidates – is this: a candidate listed first by a majority is elected; if there is no such candidate then the candidate listed first least often[†] is eliminated from all voters' lists (lifting some candidates from second to first place on some lists), and a candidate listed first by a majority of the emended lists is elected; if there is no such candidate, the procedure is repeated until a candidate with a majority is found. The method has many drawbacks: it is subject to Arrow's paradox (Borda's example above shows it since C, occupying the fewest first places, is eliminated, yet C defeats A if B is not a candidate and defeats B if A is not a candidate), it is not Condorcet-consistent, and if one or several voters change their lists by raising the winning candidate's place this could result in the election of another candidate. Moreover, it provides little to no information concerning the relative importance of the candidates and their ideas.

Well before Condorcet, Llull had proposed a more general method that is Condorcet-consistent: to rank the candidates according to their total number of wins (plus ties) in head-to-head confrontations with all other competitors. This approach differs from that of Cusanus and Borda: their method relies instead on each candidate's total number of votes in head-to-head confrontations with all other competitors. *Dasgupta-Maskin's method* [14,15] is a marriage of the two: (1) rank the candidates by their numbers of wins (thus giving precedence

[†] This arbitrary choice could be replaced by eliminating the candidate listed last most often. No reason other than "common sense" – so fallible when it comes to voting! – has been advanced for the choice made.

to the idea of Llull and Condorcet), (2) resolve any ties by using Borda's method. But it too is subject to the Arrow paradox (e.g., in the last example each candidate wins once, so recourse must be made to Borda giving the result *A>B>C;* but if *C* drops out the method elects *B*).

None of these methods overcome the Arrow paradox. Thus the natural question: is there *any* reasonable method whose inputs are voters' preferences over all pairs of candidates that does? That, of course, depends on the meaning of "reasonable." Consider it to mean the following set of "four minimal requirements":

- *Impartiality*: Voters and candidates are treated equally.
- *Unanimity*: A candidate judged to be the best by all voters is elected.
- *Decisiveness*: If the decision decrees *A* leads *B* and *B* leads *C*, then *A* leads *C*.
- *Coherence*: Adjoining or withdrawing a candidate does not change the order-of-finish.

Impartiality is the democratic ideal. Unanimity is clear. Decisiveness demands an unambiguous, transitive order-of-finish, in particular avoiding the Condorcet paradox. Coherence excludes the Arrow paradox.

Arrow's celebrated "impossibility" theorem [1] proves that when there are *at least three candidates* exactly one method satisfies the last three properties: dictatorship, namely, the preferences of exactly one voter must decide. Thus *no* method satisfies them all.

Comparing two candidates

With but two candidates there can be no cycle of preferences in the results, no intransitivity, no Arrow paradox. Does this imply that in the exceptional circumstance of two candidates, a "majority decision" that answers the question – "which candidate do you prefer?" – is acceptable?

The argument that first-past-the-post measures badly when there are several candidates is just as valid when there are but two candidates. Moreover, the fact that it is not decisive when there are at least three candidates makes it suspicious. May's [23] characterization, however, is often taken as definitive proof that majority rule is the good choice for two candidates. It proves that majority decision is the unique method that satisfies the three properties:

- *Impartiality*: Voters and candidates are treated equally.
- *Decisiveness*: Either one candidate is elected, or there is a tie.
- *Monotonicity*: If *A* is elected or in a tie with *B* and one voter changes in favor of *A* then *A* is elected.

However, May is careful to point out that the "group choice … must depend only upon individual preferences … with respect to this pair of alternatives," adding in a footnote, "the realism of this condition may be questioned." We agree with the footnote. Moreover, all methods based on comparisons fail to satisfy some desirable concept of monotonicity when

there are more than two candidates (e.g., when some voter(s) place a winner higher on their list(s) he may lose with two-past-the-post and the alternative vote; and when some voter(s) place a loser lower on their list(s) the winner may end up losing with any decisive method).

Suppose that an electorate's distribution of opinions is:

	10%	31%	15%	16%	15%	13%
A:	*Excellent*	*Excellent*	*Good*	*Good*	*Poor*	*Poor*
B:	*Good*	*Poor*	*Excellent*	*Poor*	*Excellent*	*Good*

Here, for example, 10% of the electorate rates *A Excellent* and *B Good*, so these voters may be assumed to vote for *A* and not *B* in a first-past-the-post election. The distributions of the evaluations of the candidates

	Excellent	*Good*	*Poor*
A:	41%	31%	28%
B:	30%	23%	47%

clearly show that the electorate evaluates *A* well above *B* since *A* is judged to be *Excellent* and *Good* by larger parts of the population and *Poor* by smaller parts.‡

What is the "majority decision"? When the *opinions* are distributed as above *A* is elected with 10%+31%+16%=57% of the votes. However, there are other distributions of *opinions* having the same distributions of *evaluations* that reach the opposite "majority decision":

	41%	25%	6%	5%	23%
A:	*Excellent*	*Good*	*Good*	*Poor*	*Poor*
B:	*Poor*	*Excellent*	*Poor*	*Excellent*	*Good*

Here *B* is elected with 25%+5%+23%=53% of the votes. Majority vote may well fail to elect the candidate the electorate evaluates best. There is evidence to show that this can occur in practice (see anon).

Evaluating candidates

Suppose then that voters *evaluate* candidates in a common scale of merit. The natural question that immediately arises is that posed by Arrow when voters *compare* candidates: Is there, when voters evaluate candidates, a reasonable method that chooses between every pair of candidates, i.e., that is at once impartial, unanimous, decisive, coherent *and* monotone?

The answer is yes: there are infinitely many that satisfy the four minimal requirements and are monotone (in any of the many ways in which monotonicity is defined). All must satisfy one essential requirement: *only* the candidates' distributions of evaluations may determine the

‡ In technical parlance, *A* dominates *B* stochastically.

order among them. Who gave what grade cannot be taken into account [6]. The anomaly observed in the example of the previous section is eliminated for it is not the comparative opinions of the electorate that count, only its collective evaluations. Not only does comparing deny voters adequate means to express themselves but it is the culprit that leads to Arrow's paradox, Condorcet's paradox, and still other difficulties.

The question then becomes: what method should be chosen among the infinitely many? Other desirable properties may be satisfied. Fortunately, different desiderata lead to the same solution: *majority judgment*. Two are emphasized here.

The first is to incite voters to express themselves honestly or to make it as difficult as possible for a voter to successfully manipulate by using a strategy that is not the true expression of his opinions. This is of obvious importance because the outcome of an election should be the amalgamation of voters' true opinions rather than that of calculated strategies at variance with their honest opinions. Several different criteria may be used to formulate this idea: they all lead to majority judgment. Empirical data supports the claim that majority judgment best resists strategic manipulation. Majority judgment may be characterized as the unique method that is impartial, unanimous, decisive, coherent and best resists strategic manipulation.

The second concerns measurement, for as discussed earlier, *elections measure*. The theory of measurement shows the crucial importance of distinguishing between ordinal and cardinal scales of measurement. "When measuring some attribute of a class of objects or events, we associate numbers … with objects in such a way that the properties of the attribute are faithfully represented as numerical properties" ([22], p. 1). Stars awarded by the Michelin Guide, the numbers from 0 to 10 used to measure pain, the letter combinations used by the credit rating agencies Standard & Poor's and Moody's are examples of ordinal scales: it is (in the language of measurement theory) *meaningless* to add such representations or to calculate averages of them. Days of the calendar, meters of length, degrees Celsius or Fahrenheit of heat are examples of cardinal scales: it is perfectly meaningful to add or average them.

The traditional conception of voting used in Arrow's theory has no common scale of measurement: only two-by-two comparisons have significance. Some students of voting (e.g., see [20,7,8,29]) advocate or have tested methods of election where voters assign discrete numerical values (0 to 2, or to 3, or to 10, or to 20 or to 100) and the candidates are ranked according to their total or average scores. But this is meaningless in the sense of measurement theory because for a voter to add 1 point more to a score does not mean the same thing over the entire scale: for example, in many contexts increasing a 3 to a 4 (in a scale of 0 to 20) is relatively easy to do, whereas increasing a 17 to an 18 is much harder (e.g., as in wine tasting [4]). It is not meaningful to sum such numbers.

Between these two extremes – no scale and a cardinal scale – there is a valid alternative: an ordinal scale. Majority judgment may be characterized as the unique ordinal method that is monotone and consensual [2,6]).

Majority judgment: 2012 French presidential election

Majority judgment has been tested in the last two French presidential elections (in 2007 and 2012) as well as several other elections and has been used by several juries. It is explained here in terms of a national poll conducted at the request of Terra Nova by OpinionWay in the period April 12-16, 2012 (a week or so before the official first-round on April 22). A representative sample of 993 voters participated. Its objective was to compare majority judgment with other methods of voting.

Poll's first question: "For which of the following candidates are you most likely to vote for next Sunday?" The results of this first-past-the-post poll varied a bit from the actual percentages on election day (April 22) so a set of 773 (of the 993) ballots was found for which the tallies very closely match the actual result[§] (see table 1).

Hollande	Sarkozy	Le Pen	Mélenchon	Bayrou
28.63%	27.27%	17.91%	11.00%	9.09%

Joly	Dupont-Aignan	Poutou	Arthaud	Cheminade
2.31%	1.49%	1.22%	0.68%	0.41%

Table 1: Poll, first-past-the-post results, 1st round (773 ballots).

Naturally enough the first round results elicited extensive political commentary and analyses. Headlines proclaimed:

- Enormous success of the extreme right candidate Marine Le Pen!
- Terrible deception of the leftist Jean-Luc Mélenchon who had hoped to finish third!
- Collapse of the centrist 2007 Condorcet-winner François Bayrou!
- Tight race between François Hollande and Nicolas Sarkozy in the run-off!

These measures of the relative weights of candidates and the political ideas and ideals they represent strongly influenced not only the ensuing run-off campaign that pitted the then president Sarkozy against the Socialist candidate Hollande, but will also continue to influence the unfolding political situation of Hollande's administration in the coming months and years. It is pertinent to ask: do these numbers truly reflect the reality of public opinion or are they merely such loose approximations as to be seriously misleading?

The *poll's third question* requested participants to vote with majority judgment. The ballot used is given in table 2 (the candidates are listed in their official order). Participants were asked to check one grade or appreciation in the line of each candidate. It was specified that a candidate without a check would be counted as *to Reject* (if the voter has not bothered to formulate an opinion concerning the candidate he has implicitly rejected her).

[§] The first-past-the-post and majority judgment results for all 993 participants are given in the appendix.

As President of France,
in view of all relevant considerations,
I judge, in conscience, that each of these candidates would be:

	Outstanding	*Excellent*	*Very Good*	*Good*	*Acceptable*	*Poor*	*To Reject*
Eva Joly							
Marine Le Pen							
Nicolas Sarkozy							
Jean-Luc Mélenchon							
Philippe Poutou							
Nathalie Arthaud							
Jacques Cheminade							
François Bayrou							
Nicolas Dupont-Aignan							
François Hollande							

Table 2. Poll, majority judgment ballot.

In principle, the finer the scale the better, for it enables voters to express themselves more accurately.[**] This is limited, however, by the need for a scale that is not too fine (how to distinguish between a 17 and 18 in a 0 to 20 scale?) and is common to all voters, meaning it is used and understood in the same way. The choice of a scale of seven levels accords with psychological experimentation [24] that shows most people are able to make at most seven ±2 distinctions on linear scales, as well as with voting experiments conducted to date [6,3,16]. The results are given in table 3.

	Outstanding	*Excellent*	*Very Good*	*Good*	*Acceptable*	*Poor*	*To Reject*
Joly	0.81%	2.99%	6.51%	11.80%	14.65%	24.69%	38.53%
Le Pen	5.97%	7.33%	9.50%	9.36%	13.98%	6.24%	47.63%
Sarkozy	9.63%	12.35%	16.28%	10.99%	11.13%	7.87%	31.75%
Mélenchon	5.53%	9.50%	12.89%	14.65%	17.10%	15.06%	25.37%
Poutou	0.14%	1.36%	4.48%	7.73%	12.48%	28.09%	45.73%
Arthaud	0.00%	1.36%	4.48%	7.73%	12.48%	28.09%	45.73%
Cheminade	0.41%	0.81%	2.44%	5.83%	11.67%	26.87%	51.97%
Bayrou	2.58%	9.77%	21.71%	25.24%	20.08%	11.94%	8.69%
Dupont-Aignan	0.54%	2.58%	5.97%	11.26%	20.22%	25.51%	33.92%
Hollande	12.48%	16.15%	16.42%	11.67%	14.79%	14.25%	14.24%

Table 3. Poll results, majority judgment (773 ballots).

The distributions of "grades" of candidates clearly show how the public appreciates them. Marine Le Pen, for example, has 5.97% *Outstanding*, 7.33% *Excellent,* 9.50% *Very Good*, 9.36% *Good*, 13.98% *Acceptable*, 6.24% *Poor*, and a high 47.63% *to Reject*. Her 17.9% of

[**] With 10 candidates, first-past-the-post gives the voter 11 possible expressions of opinion (to vote for one of the 10 or none); majority judgment (with 7 grades) gives him *710=284,475,249*.

the first-past-the-post votes do not have the same meaning: it is some mixture of *Outstanding*, *Very Good* and *Good*. First-past-the-post misses the public's overwhelming rejection of her candidacy. François Hollande is the candidate with the most positive evaluations: he dominates in *Outstanding*, in *Excellent* or better, and in *Very Good* or better. François Bayrou has the least negative opinions. All of this makes eminent sense to French political analysts.

One third of the voters' highest grades are *Outstanding*, one third of them are *Excellent*, one third are *Very Good* or worse; on average, a voter rejects between 3 and 4 candidates; and, fully 30% attribute their highest evaluation to two candidates or more. Similar behavior has been observed in all real uses of and experiments with majority judgment [6,5]. Such opinions are impossible to express with first-past-the-post, approval voting, or any method such as Borda's, Condorcet's or the alternative vote that asks for a rank ordering of candidates.

From the distribution of a candidate's evaluations is deduced the electorate's evaluation of her final grade, called her *majority-grade*. And from the candidates' majority-grades and the relative support for them is deduced the order-of-finish, called the *majority-ranking*. These choices are the logical consequence of a mathematical theory that has been validated in experiments and practical applications [6,3,5].

The *majority-grade* of a candidate is the grade supported by a majority against any other grade. Nicolas Sarkozy's majority-grade is *Acceptable* because a majority of 60.38%=9.63%+12.35%+16.28%+10.99%+11.13% of voters judge he merits at least *Acceptable* (implying any lower grade has only minority support) and a majority of 50.75%=11.13%+7.87%+31.75% of voters judge he merits at most *Acceptable* (implying any higher grade has only minority support). The results are given in table 4.

Majority judgment ranking	Above majority-grade *p%*	The majority-grade α±	Below majority-grade *q%*		First-past-the-post ranking	The scores
1 Hollande	45.05%	*Good* +	43.28%		1 Hollande	28.7%
2 Bayrou	34.06%	*Good* −	40.71%		2 Sarkozy	27.3%
3 Sarkozy	49.25%	*Acceptable* +	39.62%		3 Le Pen	17.9%
4 Mélenchon	42.47%	*Acceptable* +	40.43%		4 Mélenchon	11.0%
5 Dupont-Aignant	40.57%	*Poor* +	33.92%		5 Bayrou	9.1%
6 Joly	36.77%	*Poor* −	38.53%		6 Joly	2.3%
7 Poutou	26.19%	*Poor* −	45.73%		7 Dupont-Aignant	1.5%
8 Le Pen	46.13%	*Poor* −	47.63%		8 Poutou	1.2%
9 Arthaud	24.83%	*Poor* −	49.93%		9 Arthaud	0.7%
10 Cheminade	48.03%	*to Reject*	-		10 Cheminade	0.4%

Table 4. Poll, majority judgment and first-past-the-post rankings (773 ballots).

Three parameters of the distribution of each of the candidate's grades determine the *majority-ranking* of them all: α a candidate's majority-grade, *p%* the percentage of her grades above α,

and *q%* the percentage of her grades below *α*. The majority-grade is completed with a "+" when *p>q* and otherwise with a "−".

The majority-ranking is determined as follows:

(1) A candidate with a higher majority-grade is ranked above one with a lower majority-grade.

When two candidates have the same majority-grade, four sets of voters disagree: two of these sets are for higher grades for each candidate (they correspond to the candidates' *p*'s) and two are for lower grades for each candidate (they correspond to the candidates' *q*'s).

(2) When two candidates have the same majority-grade, the largest set of the four that disagree decides: if that set is for a higher grade then the corresponding candidate leads the other, if that set is for a lower grade then the corresponding candidate trails the other (which implies that a candidate with majority-grade "*α*+" leads one with the majority-grade "*α*−").

Bayrou with the majority-grade *Good* leads Sarkozy with the majority-grade *Acceptable* (by 1). Sarkozy and Mélenchon both have the majority-grade *Acceptable*, the largest set that disagrees is the 49.25% whose voters gave a higher grade to Sarkozy, so (by 2) he leads Mélenchon. Poutou and Le Pen both have the majority-grade *Poor*, the largest set that disagrees, 47.63% of the voters, gave a lower grade to Le Pen, so (by 2) she trails Poutou. Hollande has the majority-grade *Good* + and Bayrou the majority-grade *Good* −, implying the largest set that disagrees either gave Hollande a higher grade or gave Bayrou a lower grade, so in either case (by 2) Hollande leads Bayrou (as naturally indicated by the signs). Two candidates are tied only if the three parameters – *α, p% and q%* – are exactly the same. Together, a candidate's three parameters (*p%, α, q%*) is called the candidate's *majority-gauge* since it determines his/her place in the majority judgment ranking.

As is evident majority judgment (MJ) reveals a quite different distribution of public opinion – and so of the relative importance of the contending political forces. Marine Le Pen, the extreme rightist, third with first-past-the-post is but eighth with MJ in a field of ten. She only leads the two practically unknown candidates, Nathalie Arthaud and Jacques Cheminade. Why? With MJ she has few positive evaluations and many negative ones while first-past-the-post measures only the support of her partisans, completely ignoring the intensities of their support as well as the opinions of all other voters (in this case some 82% of voters). Why, then, accord her and the political ideas of her party so much attention and importance? François Bayrou, the centrist, the Condorcet-winner of the 2007 presidential election, a distant fifth with 9% of the votes with first-past-the-post, is a solid second with MJ, one of the only two candidates evaluated as *Good* by the electorate. He easily leads the incumbent President Sarkozy and Le Pen as well. Why then, accord him and the political ideas of his

party so little attention? MJ's order of finish makes sense, but reveals a political landscape very different than does first-past-the-post. Which is to be believed?

Borda, Condorcet and approval voting: 2012 French presidential election

The *poll's second question* requested participants to vote in the ten face-to-face confrontations between the five principal candidates (to ask them do so in all 45 face-to-face confrontations was too much). The results are given in table 5. Hollande beats Bayou with 51.6% of the votes and Sarkozy with 53.9% (a higher score than in that of the run-off two weeks later but in accord with other polls at the time). Hollande is the Cordorcet-winner, defeating all other candidates in head-to-head races. It happens that MV between all pairs of the five yields an unambiguous (transitive) order of finish: Hollande > Bayrou > Sarkozy > Mélenchon > Le Pen. Le Pen, notably, is easily defeated by any one of the other four: again, why give her so much attention and importance? The idea of MV as practiced is in contradiction with itself.

Borda's method (over the five candidates) gives the same result. Hollande's Borda-score is 59.5%=(51.6%+53.9%+68.5%+64.1%)/4, Bayrou's 58.7%, Sarkozy's 51.4%, Mélenchon's 45.3%, and Le Pen's 35.0%. In this case both Borda's and Condorcet's methods agree with MJ (for the five important candidates), underlining the failings of first-past-the-post.

	Hollande	Bayrou	Sarkozy	Mélenchon	Le Pen	Borda
Hollande	-------	51.6%	53.9%	68.5%	64.1%	59.5%
Bayrou	48.4%	-------	56.5%	59.4%	70.5%	58.7%
Sarkozy	46.1%	43.5%	-------	50.5%	65.7%	51.4%
Mélenchon	31.5%	40.6%	49.5%	-------	59.7%	45.3%
Le Pen	35.9%	29.5%	34.3%	40.3%	-------	35.0%

Table 5. Poll, face-to-face majority votes (773 ballots).

This failure of first-past-the-post is, of course, exactly that identified two centuries ago by Condorcet and Borda: more information should be asked of voters, voters should be given a greater freedom in the expression of their opinions. The easiest way to do this is *approval voting* (AV): the voter is asked to designate as many candidates as he wishes, the candidates are ranked according to how often they are designated. What results might one expect in this election with AV? AV elicits more information than first-past-the-post but much less than MJ, so intuitively one would guess an "intermediate" result, Bayrou gaining and Le Pen losing with respect to first-past-the-post but much less so than with MJ.

The facts support the intuition. Approval voting was tested in parallel with the first round of this election in several voting precincts of Strasbourg, Louvigny and Saint-Etienne [8]. There were 2,340 participants. The results, given in table 6, were adjusted so as to conform with the actual national results (as we did with the data of the poll), so makes a comparison possible. They show Bayrou rising to third place and Le Pen dropping to fifth, both between their first-past-the-post and MJ ranks.

Hollande	Sarkozy	Bayrou	Mélenchon	Le Pen
49.44%	40.47%	39.20%	39.07%	27.43%

Joly	Poutou	Dupont-Aignan	Arthaud	Cheminade
26.69%	13.28%	10.69%	8.35%	3.23%

Table 6. Strasbourg, Louvigny and Saint-Etienne experiment, approval voting.

These results reveal defects of approval voting that have been repeatedly observed:

- No candidate obtains a majority of AV votes (contrary to its proponents' claims [10]). The same occurred in a 2007 French presidential experiment [6] and was even more pronounced in the French presidential Orsay experiment of 2002 where the winner (Lionel Jospin) obtained but 40.5% of the votes [6].
- Whereas Borda, Condorcet and MJ place the centrist Bayrou comfortably ahead of Sarkozy, AV does not succeed to do so.
- AV yields close to a three-way tie for second place – among Sarkozy, Bayrou, and Mélenchon – whereas Borda, Condorcet and MJ all clearly distinguish between them. Close scores are often observed with AV, making AV vulnerable to strategic manipulation (as is verified empirically below).

A computer experiment was conducted on the basis of the poll's results to obtain more insight into the differences between the various methods. 10,000 randomly chosen samples of 151 ballots were chosen from the database of the 773 ballots. For each the winner was determined by each method [15]. It turns out that almost always either the centrist or the "legitimate" candidate of the left or right is elected with all methods. The results are given in table 7. Three variants of AV are included: when "approval" means at least *Excellent* (written AV ≥ *Excellent*), when it means at least *Very Good*, and when it means at least *Good*. The methods are ordered with respect to the number of times the centrist is elected.

	Left (Hollande)	Center (Bayrou)	Right (Sarkozy)	Ties
AV ≥ *Excellent*	8910	0	859	231
First-past-the-post	5715	0	3792	479
Two-past-the-post	7791	0	1415	794
AV ≥ *Very Good*	8542	19	1200	239
Majority judgment	6576	2991	430	3
Condorcet	6169	3360	471	0
Borda	4840	5007	153	0
AV ≥ *Good*	2430	6883	130	557

Table 7. Poll, French presidential election 2012: Number of wins, 10,000 sample electorates of 151 ballots drawn randomly from database of 773 ballots. Sums of lines 10,000, except second line (due to Le Pen's 14 wins).

These results are in substantial agreement with a similar computation done on the basis of the 2007 French presidential experiment conducted at Orsay [6,3] (see table 8). 10,000 randomly chosen samples of 101 ballots were chosen from a database of 501 ballots whose first round vote closely approximates the actual national first round vote.

	Left (Royal)	Center (Bayrou)	Right (Sarkozy)	Ties
AV ≥ *Excellent*	508	3	9238	251
First-past-the-post	2112	48	7824	16
Two-past-the-post	2174	764	6675	387
AV ≥ *Very Good*	1277	1316	6753	654
Majority judgment	1321	4037	4631	11
Condorcet	663	6552	1972	436
AV ≥ *Good*	380	8801	463	356
Borda	377	9592	21	10

Table 8. Orsay experiment, French presidential election 2007: Number of wins, 10,000 sample electorates of 101 ballots drawn randomly from database of 501 ballots. Sums of lines 10,000 except Condorcet's (due to 377 occurrences of the Condorcet paradox).

The results of the two experiments suggest:

- First-past-the-post systematically eliminates the centrist and accords exaggerated importance to the extremes (assuredly the case in the last three French presidential elections of 2002, 2007 and 2012).
- The methods of Condorcet and especially Borda are biased in favor of centrist candidates. This bias is less pronounced in 2012 than in 2007 because whereas in 2007 the candidate of the center (Bayrou) was the Condorcet- and Borda-winner, in 2012 it was the candidate of the left (Hollande). Moreover, in 2012 Hollande took a center-left position.
- Majority judgment is more even-handed, giving the center more chances of winning but less so than either Condorcet or Borda.
- The meaning of "approval" completely changes the results with AV. When it means at least *Excellent* or at least *Very Good* it is very biased against the center, when it means at least *Good* it is very biased for the center. And in any case, AV is more prone to ties or near ties than MJ, and so more open to manipulation.

MJ, Borda, Condorcet, and AV: 2011 French Socialist primaries

The French Socialist Party – in a historic first for France – held a primary election in 2011 to determine their candidate for the 2012 presidential election. The first round of the primary was an ideal setting to compare methods in a relatively non-polarized election with candidates and voters all of the same broad political family. Under our supervision École Polytechnique

students [16] conducted experiments in parallel with the official voting in one voting bureau of each of two communities close to Paris, Fresnes and Alfortville. After voting officially voters were asked to vote again using different systems: in Fresnes with MJ, Condorcet and Borda; in Alfortville with MJ, Condorcet and AV.

In the Fresnes experiment (table 9) majority judgment clearly distinguishes among the various candidates, in particular between Hollande and Aubry. In contrast, first-past-the-post gives the impression of a very close race between the two. Montebourg seems completely out-classed with first-past-the-post whereas majority judgment shows that he is not far behind Aubry in the appreciations of the voters.

	Above majority-grade *p%*	Majority-grade *α±*	Below majority-grade *q%*	First-past-the-post
1 François Hollande	18.2%	*Excellent* –	49.7%	35.7%
2 Martine Aubry	48.5%	*Very Good* +	20.2%	34.5%
3 Arnaud Montebourg	33.7%	*Very Good* –	39.1%	18.5%
4 Ségolène Royal	37.5%	*Good* –	38.9%	6.0%
5 Manuel Valls	36.4%	*Good* –	40.4%	5.3%
6 Jean-Michel Baylet	27.2%	*Acceptable* –	48.2%	0.0%

Table 9. Results, 2011 Socialist Primary, Fresnes: MJ and first-past-the-post (457 ballots).

It is striking, though understandable, to see the preponderance of high evaluations obtained by all the candidates compared with those used in a presidential election. A shared political outlook leads to positive appreciations, though of different nuances. With any method of voting in which the inputs compare rather than evaluate the consequence of a candidate moving up in the opinions of voters is that others move down. Not so with majority judgment: one or more candidates can improve their standing with no corresponding loss to others; indeed, all candidates can move up (due, for example, to a particularly appealing party platform backed by all). This is why monotonicity cannot be satisfied in models based on comparisons but can be satisfied in models based on evaluations.

	Aubry	Hollande	Montebourg	Royal	Valls	Baylet	Borda
1 Aubry	-	50.2 %	68.5 %	85.0 %	85,9 %	95.5 %	77.0%
2 Hollande	49.8 %	-	65.3 %	85.4 %	87,1 %	94.8 %	76.5%
3 Montebourg	31.5 %	34.7 %	-	68.3 %	69,0 %	91.8 %	59.1%
4 Royal	15.0 %	14.6 %	31.7 %	-	54,7 %	78.2 %	38.1%
5 Valls	14.1 %	12.9 %	31.0 %	45.3 %	-	78.9 %	36.4%
6 Baylet	4.5 %	5.2 %	8.2 %	21.8 %	21,1 %	-	12.2%

Table 10. Results, 2011 Socialist Primary, Fresnes: Condorcet and Borda (457 ballots).

The Borda and Condorcet rankings agree in Fresnes, Aubry leading Hollande by a very narrow margin. With majority judgment, however, Hollande is the clear winner. This illustrates in practice what was seen could occur in theory: the candidate a majority of electors *evaluate* highest is not necessarily the candidate that wins by a majority in *comparisons*. The intuitive explanation is clear: electing a candidate on the basis of comparisons ignores the intensity of feelings for and against him, so the "loser" according to simple majority voting may in fact be more consensual. This is what happened here. The same underlying phenomenon is observable in France's 2012 presidential run-off between Hollande and Sarkozy. Hollande won a close election with 51.6% of the votes to Sarkozy's 48.4%, yet Hollande almost certainly dominated clearly in evaluations: majority judgment more surely elects the consensual candidate than simple majority voting.

Voters do not always vote sincerely in accord with their preferences: some try to "game the vote" [25], anticipating the votes of others (from polls, media coverage and commentary) to try to make their votes really useful in view of the expected outcome. A good method should incite honest voting and resist successful manipulation. Take a candidate whose majority-grade is (say) α = *Good,* with p% of the voters assigning him a higher grade and q% a lower grade, and consider a set of voters who accorded him (say) *Very Good*. Presumably they are disappointed, believing the candidate should have a higher majority-grade, but with MJ they are unable to change α, p% *or* q% by raising the grade they assign to *Outstanding* or *Excellent*. Similarly, a set of voters who accorded the candidate a lower grade (*Acceptable*, *Poor* or *to Reject*) – so who are disappointed because they believe he should have a lower majority-grade – cannot change the three parameters by lowering the grade they assign. Thus, if a voter wishes for a candidate to be accorded a particular majority-grade then his best strategy is to assign him precisely that grade: honesty is the best policy in grading. There are as well other criteria by which it may be proven that majority judgment is as robust as possible against strategic manipulation [6]. It is impossible to resist successful manipulation entirely with any method, but MJ was chosen as that method which best resists.

Some skeptics question the use of majority judgment in a field of two candidates, saying majority decision is the only acceptable rule and suggesting that with MJ a voter would always give the highest possible grade to his preferred candidate and the lowest possible grade to her opponent. If voters actually behaved in this way then the MJ outcome would simply be the usual majority decision outcome (so why object?). But such behavior assumes that the only relevant outcome to a voter participating in a two-candidate election is who wins. How the candidates split the vote may also be of importance to a voter, and so could influence the (calculated) decision on how to vote and whether to vote at all. Moreover, most voters, we believe, prefer to express their honest opinions when posed a question. A national poll conducted on the day of the first-round of the 2007 French presidential election supports this belief. It asked voters which of seven factors most determined their votes. The previous 2002 election had been profoundly shocking: the extreme right candidate Jean-Marie Le Pen had reached the second round, denying the expected run-off between Chirac and Jospin. So there was an accrued collective realization that honest voting could well lead to another catastrophe. Nevertheless, only 21% of the voters affirmed that a deliberate strategic vote

different from their preference was the determining explanation for their votes. Look at the question from your personal point of view. Suppose you – a voter – are confronted with two candidates. Often (in this 21st century) you don't much care for either; you may, sometimes, like both equally; or you may support aspects of both candidates' ideas and persons and oppose others. Would you, in such circumstances, give one candidate the highest grade, the other the lowest? That seems questionable.

Manipulability was tested too. 10,000 random samples of 101 (and of 151) ballots were chosen of the 457 in Fresnes. In each sample, the set of voters who preferred the runner-up to the winner was identified: those ballots were changed by giving, with a probability of 30%, the highest grade or the highest place in the rank-order (depending on the method used) to the runner-up and the lowest grade or the lowest place to the winner. When the result changes the manipulation is successful, otherwise it is not. The percentages of successful manipulations are given in table 11. As many other experiments have shown, Borda's method is extremely manipulable. Of course, with any method, the rate of successful manipulations will augment as the margin of victory diminishes. But majority judgment better resists manipulation than any other method, as is confirmed in other similar empirical tests [6,5].

	101 ballots	151 ballots
Majority judgment	57.3%	69.2%
Borda	99.9%	100%

Table 11. Successful manipulations, 2011 Socialist Primary, Fresnes (457 ballots).

The results of the Alfortville experiment are given in table 12. Majority judgment again clearly distinguishes between the candidates whereas approval voting raises doubts: (1) all the socialist candidates are "approved" by comfortable majorities (Baylet is a member of another party), three with huge majorities; and (2) 85% and 87% "approval" give Aubry and Hollande essentially the same legitimacy, although the finer measure of the electorate's collective evaluation provided by MJ's majority-gauges (*p%, α±, q%)* shows Hollande dominates comfortably.

	Above majority-grade *p%*	Majority-grade *α±*	Below majority-grade *q%*	First-past-the-post	Approval voting
1 Hollande	40.1 %	*Very Good* +	25.4%	37.7%	87.3%
2 Aubry	33.1%	*Very Good* +	30.6%	29.2%	85.2%
3 Montebourg	39.8%	*Good* +	36.3%	12.5%	64.1%
4 Valls	28.5%	*Good* –	44.7%	10.0%	53.2%
5 Royal	27.1%	*Good* –	47.2%	10.3%	53.5%
6 Baylet	41.5%	*Poor* +	28.9%	0.4%	25.7%

Table 12. Results, 2011 Socialist Primary, Alfortville (292 ballots).

The empirical tests to assess the rates of successful manipulation in this experiment are given in table 13 [15]. It is difficult to change the result with majority judgment, but Hollande was well in the lead. Approval voting, however, is vulnerable to strategic manipulation.

	101 ballots	151 ballots
Majority judgment	20.8%	15.0%
Approval voting	82.0%	90.8%

Table 13. Successful manipulations, 2011 Socialist Primary, Alfortville (292 bulletins).

Conclusion

Many of the scientific community have steadfastly stayed with "the legitimating force of the majority rule," albeit interpreted differently, so leading to the support of a number of different methods based on *comparing*: Condorcet's, Borda's, a hybrid of the two, approval voting, the alternative vote. Many too have resisted the idea of *evaluating the merit* of candidates. This may be due, in part, to confusion between measuring *merit* – as is done with majority judgment – and measuring the *utility* of outcomes to voters (e.g., [28,20]). The utility to a voter is a measure of his satisfaction with the outcome of an election. It is a relative notion that concerns the results or "outputs" of an election, whereas the evaluation of the merits of candidates concerns the "inputs" of an election. Thus, for example, a left-leaning voter was undoubtedly very satisfied to see Chirac crush Le Pen in the election of 2002. But this same voter would have been dissatisfied had Chirac defeated Jospin in the second-round of that election. This has nothing to do with his evaluations: in both cases he would likely have judged Chirac as *Acceptable* or *Poor*, Jospin as *Good* or better, and Le Pen as *to Reject*.

The scientific community concurs that usual simple majority decision – first-past-the-post – is to be eschewed when there are at least three candidates; it concurs also that simple majority decision should be the choice when there are but two candidates. Indeed, all of the advocated methods based on comparisons become simple majority decision when there are but two candidates. And yet, we believe and argue, majority decision between two may fail, and has. Jacques Chirac defeated Jean-Marie Le Pen with a crushing 82% of the vote in the run-off of 2002, but that percentage in no way measured his support in the nation. Close scores between two candidates may mean little as well: voters may like both candidates, may like neither, or may like one and not the other: these differences can be captured only with a finer measure of voters' opinions. Moreover, close scores occur sufficiently often to be important: witness the 2000 U.S. presidential race that was decided by an official margin of 537 votes (giving Bush all of Florida's 25 Electoral College votes), the 2012 contest between François Fillon and Jean-François Copé for the presidency of the UMP party in France that was decided by a margin of 98 votes, or the 2008 contest for the head of the Socialist party in France won by Martine Aubry with 67,451 votes to Ségolène Royal's 67,349, a margin of 102 votes.

Condorcet's method is not decisive – it may lead to no winner – and suffers from the defects of simple majority decision between two candidates.

Borda's method is decisive, but it is open to Arrow's paradox, meaningless in the sense of measurement theory, highly manipulable, and unduly biased in favor of centrists.

Dasgupta-Maskin's method is a hybrid of Condorcet's (or Llull's) and Borda's. Based on comparisons – so deficient for the reasons developed above – it seeks to satisfy the four minimal requirements and to resist manipulation. To prove it satisfies these requirements, instead of enlarging voters' possible expressions of opinion, it must restrain them (implicitly assuming that voters' rank-order inputs are naturally expressible along a clear-cut ideological ordering of the candidates and the parties they represent that is held in common by all). There is ample experimental evidence that shows voters do not behave in accord with this restriction.[††] In any case, it inherits the bad properties of Borda's method and, in particular, admits the Arrow paradox, as has been observed in figure skating competitions [5] and in the famous 1976 wine competition known as the "judgment of Paris" [4]. Known in the figure skating world as the OBO (One-by-One) system – introduced for figure skating in 1998 – it was used in only one winter Olympics (that of 2002) before being definitively discarded for a system based on points (and not comparisons).

The yes/no or two levels of approval voting are insufficient to judge candidates. In polarized elections pitting different ideologies and policies against each other such as a presidential race, often no candidate receives a majority of "approvals" and the results are not sufficiently different to impart legitimacy to the winner. On the other hand, in a non-polarized election such as a primary, often many candidates end with strong "majorities," so again none has a really clear and distinctive mandate. In both cases experience shows that scores may be close, so the results are all the more manipulable.

It is sometimes claimed that approval voting is Condorcet-consistent and resists manipulation (see e.g., [10]). It is important to understand exactly what this means. If voters express themselves honestly, approval voting is not Condorcet-consistent even when there are only two candidates, as the following example shows:

	20%	15%	20%	30%	15%
A	*Excellent*	*Very Good*	*Acceptable*	*Acceptable*	*Poor*
B	*Very Good*	*Good*	*Good*	*Poor*	*to Reject*

80% of the voters prefer *A* to *B*. However, if "approve" means at least *Good* (as is often suggested by approval voting advocates) then *B* wins with 55% "approve" to *A*'s 35% "approve." The fact is that approval voting is sure to elect a Condorcet-winner *only* when all the voters behave strategically and they care only about who wins; moreover, this is provable only for some concepts of what constitutes a solution in an abstract election game.

[††] For example, the 1999 presidential election of the Social Choice and Welfare Society [11,26]. Or the approval voting experiment conducted in parallel with the 2002 French presidential election in Orsay. 16 candidates ran, so the number of different possible ballots consistent with single-peakedness is 137; however, the actual total number of different ballots was 813 (of 2,587 valid ballots) [6].

The alternative vote – or instant run-off voting (IRV) – is decisive, but it is opaque, difficult for voters, admits the Arrow paradox, and not "monotone" in the sense that more support for a winner may make him lose. In Australia, where it is practiced, there may be a dozen candidates; listing them all is so difficult that there is a provision that permits a voter to opt for a party's specification of the ordering – and that, of course, is arrived at on the basis of wheels and deals among parties.

In contrast with the theoreticians of voting, those who judge competitions – of wines, figure skaters, gymnasts, divers, musicians, etc. – have increasingly abandoned methods based on comparisons to methods based on evaluations of merit. Typically the evaluations are in a scale of numbers that are carefully defined and well understood by the judges. They are not, however, cardinal measures because 1 point more at the low end of the scale means something altogether different than 1 point more at the top end of the scale. Nevertheless, the order-of-finish of the competitors is usually determined by the sums of their scores, which is meaningless in the sense of measurement theory. Moreover, such methods are the most manipulable of all. The decisions they render are assuredly not majority decisions for any one judge can change the final ranking – sometimes to his entire satisfaction – by changing the scores he assigns to competitors (up and down) [4,5].

The practice in judging competitions has taken an important first step: instead of comparing it evaluates. A second step needs to be taken too: to change the procedures for amalgamating the different evaluations of judges into the jury's final evaluation of each competitor. Some have taken a small step in this direction by eliminating the top and bottom grades or the top two and bottom two grades [6]. Terra Nova – "an independent progressive think tank whose goal is to produce and diffuse innovative political solutions in France and Europe" – has taken both steps in voting: it has included majority judgment in its recommendations for reforming the presidential election system of France [30].

References

[1] K. Arrow. 1951 (2nd edition 1963). *Social Choice and Individual Values*. New Haven CT: Yale University Press.

[2] M. Balinski and R. Laraki. 2007. "A theory of measuring, electing and ranking." *Proceedings of the National Academy of Sciences, USA.* 104 872-8725.

[3] M. Balinski and R. Laraki. 2011. "Election by majority judgment: experimental evidence." In B. Dolez, B. Grofman, and A. Laurent (eds.), *In Situ and Laboratory Experiments on Electoral Law Reform: French Presidential Elections* 13-54. Berlin: Springer.

[4] M. Balinski and R. Laraki. 2012. "How best to rank wines: majority judgment." To appear in E. Giraud-Héraud and M.-C. Pichery (eds.), *Wine Economics: Quantitative Studies and empirical Observations* (provisional title). London: Palgrave.

[5] M. Balinski and R. Laraki. 2013. "Judge: don't vote!"

[6] M. Balinski and R. Laraki. 2010. *Majority Judgment: Measuring, Ranking, and Electing.* Cambridge, MA, USA: MIT Press.

[7] A. Baujard and H.Igersheim. 2007. *Expérimentation du vote par note et du vote par approbation lors des élections présidentielles françaises du 22 avril 2007. Rapport final*. Rapport et documents du Centre d'Analyse Stratégique, Paris.

[8] A. Baujard, F. Gavrel, H.Igersheim, J.-F. Laslier and I. Lebon. 2013. "Vote par approbation, vote par note. Une experimentation lors de l'élection présidentielle du 22 avril 2012." *Revue Economique* (to appear).

[9] Le Chevalier de Borda. 1784. "Mémoire sur les élections au scrutin." *Histoire de l'Académie royale des sciences* 657-665.

[10] S. Brams and P. Fishburn. 1983. *Approval Voting*. Boston: Birkhäuser.

[11] S. Brams and P. Fishburn. 2001. "A nail-biting election." *Social Choice and Welfare* 18 409-414.

[12] Le Marquis de Condorcet. 1785. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix***.** Paris: l'Imprimerie royale.

[13] A. Copeland. 1951. "A 'reasonable' social welfare function." Seminar on Applications of Mathematics to the Social Sciences. University of Michigan, Ann Arbor.

[14] P. Dasgupta and E. Maskin. 2004. "The fairest vote of all." *Scientific American* 290, March, 92-97.

[15] P. Dasgupta and E. Maskin. 2008. "On the robustness of majority rule." *Journal of the European Economics Association* 6 949-973.

[16] B. Faveau, J. Gonzalez-Suitt, A. Guyon, T. Hennion, X. Starkloff, and S. Thibault. 2012. "Vers un système du vote plus juste?" Group project, Ecole Polytechnique.

[17] C. Girard. 2010. "Acclamation voting in Sparta : an early use of approval voting." In J.-F Laslier et R. Sanver (eds.), *Handbook of Approval Voting*. Berlin: Springer, pp. 15-18.

[18] G. Hägele and F. Pukelsheim. 2001. "Llull's writings on electoral systems." *Studia Lulliana* 41 3-38.

[19] G. Hägele and F. Pukelsheim. 2008. "The electoral systems of Nicolas Cusa in the Catholic Concordance and beyond." In ed. by G. Christianson et al, *The Church, the Councils and Reform: Lessons From the Fifteenth Century,* 229-249. Washington, DC: Catholic University Press of America.

[20] C. Hillinger. 2004. "Utilitarian collective choice and voting." Discussion paper 2004-25, Department of Economics, University of Munich.

[21] P. Kurrild-Klitgaard. 1999. "An empirical example of the Condorcet paradox of voting in a large electorate." *Public Choice* 107 1231-1244.

[22] D. H. Krantz, R.D. Luce, P. Suppes, and A. Tversky. 1971. *Foundations of Measurement*. Vol. 1. New York: Academic Press.

[23] K. O. May. 1952. "A set of independent necessary and sufficient conditions for simple majority decision." *Econometrica* 20 680-684.

[24] G. Miller. 1956. "The magical number seven, plus or minus two: some limits on our capacity for processing information." *Psychologial Review* 63 81-97.

[25] W. Poundstone. 2008. *Gaming the Vote: Why Elections Aren't Fair*. New York: Hill and Wang.

[26] D. G. Saari. 2001. "Analyzing a nail-biting election." *Social Choice and Welfare* 18 415-430.

[27] Wojciech Sadurski. 2008. "Legitimacy, political equality, and majority rule." *Ratio Juris* 21 39-65.

[28] A. Sen. 1970. *Collective Choice and Social Welfare*. San Francisco: Holden-Day.

[29] W. Smith. 2007. RangeVoting.org http://rangevoting.org/

[30] Terra Nova [2011]. *Rendre les élections aux électeurs : le jugement majoritaire.* April 21. <http://www.tnova.fr/note/rendre-les-lections-aux-lecteurs-le-jugement-majoritaire>

Appendix

The results of the OpinionWay poll conducted April 12-16, 2012 including all 993 participants (a representative sample of the French electorate):

	Outstanding	***Excellent***	***Very Good***	***Good***	***Acceptable***	***Poor***	***To Reject***
Joly	0.81%	2.72%	5.94%	11.18%	16.01%	25.88%	37.46%
Le Pen	4.43%	5.74%	8.36%	9.26%	13.80%	8.36%	50.05%
Sarkozy	8.96%	11.18%	15.31%	11.48%	12.79%	8.96%	31.32%
Mélenchon	4.93%	8.46%	11.88%	14.30%	18.03%	16.01%	26.38%
Poutou	0.30%	1.31%	4.13%	7.75%	12.79%	28.30%	45.42%
Arthaud	0.10%	1.11%	3.32%	6.55%	13.80%	26.08%	49.04%
Cheminade	0.30%	0.60%	2.01%	5.44%	11.88%	27.19%	52.57%
Bayrou	3.22%	8.96%	20.14%	24.77%	20.44%	12.49%	9.97%
Dupont-Aignan	0.40%	1.91%	5.34%	11.68%	18.93%	26.28%	35.45%
Hollande	10.27%	14.10%	15.61%	14.30%	16.41%	14.20%	15.11%

Majority judgment ranking	Above majority-grade *p%*	The majority-grade α±	Below majority-grade *q%*		First-past-the-post ranking	The scores
1 Bayrou	32.33%	*Good* –	42.90%		1 Sarkozy	28.03%
2 Hollande	39.98%	*Good* –	45.72%		2 Hollande	26.91%
3 Sarkozy	46.93%	*Acceptable* +	40.28%		3 Le Pen	14.80%
4 Mélenchon	39.58%	*Acceptable* +	42.40%		4 Bayrou	12.00%
5 Dupont-Aignan	38.27%	*Poor* +	35.45%		5 Mélenchon	11.32%
6 Joly	36.66%	*Poor* –	37.46%		6 Joly	2,80%
7 Poutou	26.28%	*Poor* –	45.42%		7 Poutou	2.02%
8 Arthaud	24.87%	*Poor* –	49.04%		8 Dupont-Aignan	1.23%
9 Le Pen	49.95%	*to Reject*	--		9 Arthaud	0.56%

10 Cheminade	47.43%	*to Reject*	--		10 Cheminade	0.34%

These results differ from those with 773 participants (whose first-round votes closely matched the true national percentages). Bayrou is the close winner over Hollande, and Le Pen – *to Reject* – falls to ninth place from eighth.

The numbers of wins with different methods based on 10,000 random samples of 151 ballots from the 993 of the poll[‡‡]:

	Left (Hollande)	Center (Bayrou)	Right (Sarkozy)	Ties
AV ≥ *Excellent*	8010	0	1626	363
First-past-the-post	3967	0	5564	467
Two-past-the-post	7929	1	1347	720
AV ≥ *Very Good*	7633	96	1921	350
Majority judgment	5023	4672	302	2
Condorcet	5220	4452	327	0
Borda	3740	6188	72	0
AV ≥ *Good*	2109	7301	94	495

The successful manipulations in the Terra Nova poll:

	Database of all 993 ballots		Database of 773 ballots	
	101 ballots	151 ballots	101 ballots	151 ballots
Condorcet	33.3%	36.0%	31.7%	34.2%
Majority judgment	50.9%	58.0%	42.6%	46.6%
First-past-the-post	51.9%	63.7%	47.1%	58.6%
Borda	84.4%	90.4%	84.5%	90.9%

‡‡ The sums in each line are 10,000, except that: Mélenchon wins 1 time with AV ≥ *Excellent,* Majority judgment, Condorcet and AV ≥ *Good*; he wins 3 times with Two-past-the-post; and Le Pen wins 2 times with First-past-the-post.