ECOLE POLYTECHNIQUE
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

MAJORITY MEASURES

Michel BALINSKI
Rida LARAKI

March 2015

Cahier n° 2015-05

DEPARTEMENT D'ECONOMIE
Route de Saclay
91128 PALAISEAU CEDEX
(33) 1 69333033
http://www.economie.polytechnique.edu/
mailto:chantal.poujouly@polytechnique.edu

Majority Measures*†

Michel Balinski
CNRS, École Polytechnique, Paris, France

and

Rida Laraki
CNRS, LAMSADE, Université Paris-Dauphine and
École Polytechnique, Paris, France

"But let them measure us by what they will;
We'll measure them a measure, and be gone."
— William Shakespeare, *Romeo and Juliet*

Abstract

The validity of majority rule in an election with but two candidates—and of Condorcet consistency—is challenged. Axioms based on measures—paralleling those of K. O. May characterizing majority rule for two candidates that are based on comparisons—lead to another method. It is unique in agreeing with the majority rule when the electorate is "polarized" and meets R. A. Dahl's requirement that an apathetic majority not defeat an intense minority. It accommodates any number of candidates and avoids both the Condorcet and Arrow paradoxes.

Key words: measuring, ranking, electing, majority rule, Condorcet consistency, tyranny of majority, intensity problem, majority judgment, majority-gauge.

JEL classification: D71

*This work was supported in part by a grant administered by the French National Research Agency as part of the Investissements d'Avenir program (Idex [Grant Agreement No. ANR-11-IDEX-0003-02/Labex ECODEC No. ANR11- LABEX-0047]).

†The authors are indebted to Paul Edelman, Hervé Moulin, and especially Jack Nagel for very helpful comments and suggestions.

Introduction

Methods of voting *measure* the support given by an electorate (or jury) to each of several competitors to determine an order of finish and a winner among them.

Surprisingly, no traditional method nor any in use contains a hint about how to measure the support of *one* candidate (though in a significant number of elections, including U.S. congressional elections, there is but one candidate). Instead, the theory of voting (or of social choice) elevates to a basic distinguishing axiom the faith that in an election between *two* candidates majority rule—based on comparing candidates—is the only proper rule. Every traditional method of voting ever proposed is claimed to reduce to the majority rule when there are only two candidates. As is well known, comparing candidates when *three or more* compete leads to insurmountable paradoxes—notably, Condorcet's and Arrow's—and so to an inconsistent theory and sometimes dubious decisions.

Thus the only situation where the majority rule can affirm anything unambiguously is when there are exactly two candidates: it says nothing when there is only one, it is incoherent when there are three or more. Everyone seems to be convinced that using the majority rule to choose one of two candidates is infallible: since infancy who in this world has not participated in raising their hands to reach a collective decision? Tocqueville believed, "It is the very essence of democratic governments that the dominance of the majority be absolute; for other than the majority, in democracies, there is nothing that resists" ([23], p. 379)[1].

Every student of social choice seems to accept K. O. May's axiomatic justification of it, and much of the literature on the theory of voting takes *Condorcet consistency*—that assures the election of a candidate who defeats each of the others separately in majority votes—to be either axiomatic or a most desirable property. R. A. Dahl (who also cited Tocqueville) stated in 1956:

> "The only rule compatible with decision-making in a populistic democracy is the majority principle . . . [which] prescribes that in choosing among alternatives, the alternative preferred by the greater number is selected. That is, given two or more alternatives x, y, etc., in order for x to be government policy it is a necessary and sufficient condition that the number who prefer x to any alternative is greater than the number who prefer any single alternative to x." ([9], pp.37-38)

That view seems to have gone unchallenged, though Dahl himself raised questions about it due to the realistic existence of the intensities of preferences of voters.

Beginning with very simple concepts that are basic to the meaning of an electorate's global measure of *one* candidate, this article shows how majority rule can easily go wrong when voting on but *two* candidates, let alone more.

[1] Our translation of: "Il est de l'essence même des governments démocratiques que l'empire de la majorité y soit absolu; car en dehors de la majorité, dans les démocraties, il n'y a rein qui résiste."

In consequence it seems unreasonable to give so much importance to Condorcet consistency. The article goes on to describe a consistent approach for any number of candidates.

Background

The *majority rule* (MR) in a field of two elects that candidate preferred to the other by a majority of the electorate.

A rule provides an outcome: one candidate is the winner against the other, there is a tie between them, or the rule is incomplete and says nothing. May [17] proved that the majority rule is the one rule that satisfies the following five simple properties in an election with two candidates.

Axiom* 1 (Based on comparisons) *A voter expresses her opinion by preferring one candidate or being indifferent between them.*[2]

Axiom* 2 (Anonymous) *Interchanging the names of voters does not change the outcome.*

Axiom* 3 (Neutral) *Interchanging the names of the candidates does not change the outcome.*

Anonymity stipulates equity among voters, neutrality the equitable treatment of candidates.

Axiom* 4 (Monotone) *If candidate A wins or is in a tie with the other and one or more voters change their preferences in favor of A then A wins.*

A voter's change in favor of A means changing from a preference for B to either indifference or a preference for A, or from indifference to a preference for A.

Axiom* 5 (Complete) *The rule guarantees an outcome: one of the two candidates wins or they are tied.*

The underlying assumption of traditional voting theory is that each voter has preferences expressed as comparisons: for her one candidate is either better or worse than another, or she is indifferent between them. With three or more candidates voter rationality implies that a voter's preferences may be expressed as a rank-ordering of the candidates, a list going from the most to the least preferred candidate (perhaps with some indifferences).

When there are at least three candidates two paradoxes raise their ugly heads. The *Condorcet paradox* [8] shows that an electorate can make a candidate A the MR-winner against candidate B, B the MR-winner against candidate C, and C the MR-winner against A. It has been observed in elections [16] as well as in wine-tasting ([2], pp. 156-159) and figure skating [4].

[2] In a footnote May had the wisdom to admit, "The realism of this condition may by questioned."

How to generalize the majority rule to more than two candidates has been the focus of much of the theory of social choice (e.g., [14, 27, 26, 10]). Many of today's electoral systems do so by asking voters to vote for at most one candidate and declaring the winner to be the candidate with the most votes; or, in some systems, by going on to a run-off between the top two vote-getters if neither has an absolute majority in the first round. Call these systems MR^+. The *Arrow paradox* [1] shows that with identical preferences an electorate can make candidate A the MR^+-winner but makes B the MR^+-winner against A if some other candidate C withdraws. It has occurred frequently [3]. Al Gore would almost certainly have carried the state of Florida in the 2000 U.S. presidential race—and so would have been elected president—had Ralph Nader not been a candidate. In the 2007 French presidential election François Bayrou finished behind Ségolène Royal and Nicolas Sarkozy in the first-round yet all the polls showed he would easily have defeated either Royal or Sarkozy in head-to-head encounters.

One candidate

> "It is clear that every measurement—or rather every claim of measurability—must ultimately be based on some immediate sensation . . . In the case of utility the immediate sensation of preference—of one object or aggregate of objects against another—provides this basis. . . . All this is strongly reminiscent of the conditions existent at the beginning of the theory of heat: that too was based on the intuitively clear concept of one body feeling warmer than another, yet there was no immediate way to express significantly by how much, or how many times, or in what sense." ([20], p. 16)

So began von Neumann and Morgenstern's argumentation for a theory of measurable utility. Voting and the traditional theory of social choice has adhered to comparisons, a voter "feeling" one candidate better than another but denied the possibility of expressing his feelings more precisely.

However, the practice in virtually every other instance that ranks entities is to evaluate each of them (see [2], chapters 7 and 8). The *Guide Michelin* uses stars to rate restaurants and hotels. Competitive diving, figure skating, and gymnastics use carefully defined number scales. Wine competitions use words (*Excellent, Very Good, Good, Passable, Inadequate, Mediocre, Bad*) to which are attached numbers. Pain uses sentences to describe each element of a scale that is numbered from 0 ("Pain free") to 10 ("Unconscious. Pain makes you pass out."), a 7 defined by "Makes it difficult to concentrate, interferes with sleep. You can still function with effort. Strong painkillers are only partially effective."

In the political sphere polls sometimes ask participants whether they *Approve* or *Disapprove* the performance of an office holder. A more probing example is a Harris poll that asked, ". . . [H]ow would you rate the overall job that President

Barack Obama is doing on the economy?" Among the answers spanning 2009 to 2014 were those given in Table 1.

	Excellent	*Pretty good*	*Only fair*	*Poor*
March 2009	13%	34%	30%	23%
March 2011	5%	28%	29%	38%
March 2013	6%	27%	26%	41%

Table 1. Measures evaluating the performance of Obama on the economy [13].

In sum it is not only natural to use measures to evaluate a performance, restaurant, wine, or politician, but necessary. To be able to measure the support a candidate enjoys a voter must be given the means to express herself. To assure that voters are treated equally, they must be confined to a set of expressions that is shared by all. To allow for meaningful gradations—different shades ranging from very positive, through mediocre, to very negative—the gradations must faithfully represent the possible expressions. As was seen such ordered evaluations are common in every day life. Call such a set of commonly held, ordered evaluations a *scale of grades.*

Axiom 1 (Based on measures) *A voter's opinion is expressed by evaluating each candidate in a scale of grades* Γ.

An electorate's *opinion profile* on the candidate is the entire set of her, his, or its grades $\gamma = (\gamma_1, \dots \gamma_n)$, where $\gamma_j \in \Gamma$ is voter j's evaluation of the candidate. It may be described as the set of grades, or the number of times each grade occurs, or the percentages of the grades' appearances (as in the Harris poll).

An aggregation function G associates a *global measure* $G(\gamma) \in \Gamma^*$ with any opinion profile. Underlying the data such as that of the Harris poll is the assumption that every voter's evaluation counts the same.

Axiom 2 (Anonymous) *Permuting the grades of any two voters does not change the electorate's global measure.*

Proposition *An anonymous global measure* $G(\lambda)$ *depends only on the set (or the distribution) of the grades* λ.

Only the grades count: which voter gave what grade has no impact on the electorate's global measure of a candidate. The most complete possible global measure is the set of grades λ itself, the number of times each grade occurs or their percentages (as in the Harris poll given in Table 1). However, in many applications the global measure takes on a simpler form. Most often, when the γ_j's are integers, their average (typically not an integer) is used as a global measure. Another possible choice, well defined for any ordered set of grades, is the majority decision γ^*: a majority evaluates the candidate as γ^* or better and a majority evaluates her as γ^* or worse (thus γ^* is the median of her grades). In this case $\Gamma^* = \Gamma$.

How are two global measures of Obama's performance on the economy at different dates to be compared? Had the March 2011 or 2013 answer been identical to that on March 2009 the performance would have been judged to be the same. In fact Obama's March 2009 evaluation dominates those of the same month in 2011 and 2013 and so is clearly better. In general, the candidate's global measure at time t, λ^t, *dominates* his global measure at time s, λ^s, when λ^t has at least as many of the highest grade as λ^s, at least as many of the two highest grades as $\lambda^s, \ldots$, at least as many of the k highest grades as λ^s, for all k, and at least one "at least" is "more." When λ^t and λ^s are both the complete set of grades (with each grade λ_j^t and λ_j^s repeated the number of times it appears) and are listed from the highest to the lowest, λ^t dominates λ^s when $\lambda_j^t \succeq \lambda_j^s$ for every j and at least one $\succeq$ is strict.

Any reasonable method of ranking should *respect domination*: namely, evaluate the candidate's performance at time t above that at time s when his global measure at time t dominates that at time s. Surprisingly, while some methods respect domination many do not.

Two candidates

How should the global measures of two candidates A and B to be compared? In the same manner that the global measures of one candidate at two different times are compared.

The basic input is an electorate's *opinion profile*: it gives the grades assigned to each candidate by every voter and may be represented as a matrix of two rows (one for each candidate) and n columns (one for each voter) displaying them. The *preference profile* of the traditional theory—voters' rank-orderings of the candidates—may be deduced from the opinion profile.

A *method of ranking* $\succeq$ is a non symmetric binary relation on candidates that associates to each opinion profile a comparison between them, $A \succeq B$ meaning that A is either better or equal than B, $A \approx B$ that they are evaluated equally, and $A \succ B$ that A is strictly better. A priori $\succeq$ is not complete, so two candidates may not be comparable. The majority rule is an example of a method of ranking that is complete.

A *point-summing method* chooses an ordinal scale—words or descriptive phrases (sometimes numbers)—and assigns to each a numerical grade. There are, of course, infinitely many ways to assign such number grades. Every voter evaluates each candidate in that scale and the candidates are ranked according to the averages of their grades. An example is the Danish educational system that has six grades with numbers attached to each[3]: *Outstanding* 12, *Excellent* 10, *Good*, 7, *Fair* 4, *Adequate* 2 and *Inadequate* 0 [25]. Any point-summing method clearly clearly respects domination. Majority judgment [2, 3, 4] is another method that respects domination. Majority rule, however, is an example of a method that does not.

[3]Why these numbers is explained anon.

The majority rule may well rank B above A when A's evaluation dominates B's, as shown in the example of Table 2.

	Very Good	*Acceptable*	*Poor*
A:	40%	36%	24%
B:	36%	34%	30%

Table 2. Example where A's grades dominate B's.

A's grades clearly dominate B's. Yet these distributions could correspond to the opinion profile:

	10%	30%	36%	24%
A:	*Very Good*	*Very Good*	*Acceptable*	*Poor*
B:	*Acceptable*	*Poor*	*Very Good*	*Acceptable*

(where, e.g., 10% of the voters evaluate A to be *Very Good* and B to be *Acceptable*). B wins with a majority of $36\% + 24\% = 60\%$.

Other opinion profiles with the same distributions of grades can give the opposite majority decision, e.g.,

	30%	10%	16%	20%	20%	4%
A:	*Very Good*	*Very Good*	*Acceptable*	*Acceptable*	*Poor*	*Poor*
B:	*Acceptable*	*Poor*	*Very Good*	*Poor*	*Very Good*	*Acceptable*

Here A wins with a majority of $30\% + 10\% + 20\% = 60\%$.

The first scenario that elects the wrong candidate is not an isolated phenomenon. The probability with which A or B wins with the majority rule is easily estimated. The set of all possible opinion profiles consistent with the electorate's assessments when no voter assigns the same grade to both candidates is given in Table 3 for values of δ in the interval $10\% \leq \delta \leq 34\%$.

		B: *Very Good*	B: *Acceptable*	B: *Poor*
	Very Good		$\delta\%$	$(40 - \delta)\%$
A:	*Acceptable*	$(46 - \delta)\%$		$(\delta - 10)\%$
	Poor	$(\delta - 10)\%$	$(34 - \delta)\%$	

Table 3. All possible opinion profiles, $10\% \leq \delta \leq 34\%$.

(here, for example, $\delta\%$ evaluate A to be *Very Good* and B to be *Acceptable*, $(34 - \delta)\%$ evaluate A to be *Poor* and B to be *Acceptable*).

When $\delta = 20\%$ the majority rule produces a tie; when $10\% \leq \delta < 20\%$ the majority rule winner is B; and when $20\% < \delta \leq 34\%$ the majority rule winner is A. Assuming all of the opinion profiles are equally likely the calculation shows A is the winner $\frac{7}{12}$th of the time, B is the winner $\frac{5}{12}$th of the time.

This simple example shows how very badly majority rule may fail when there are only two candidates: a 41.66% error rate when the outcome is crystal clear seems very high.

Any number of candidates

How should any number of candidates be compared? With but two candidates A should clearly defeat B when A's grades dominate B's. And when the grades of the performance of one candidate dominates his performance at another time the first should clearly be judged better (as for instance in comparing Obama's performances in 2009 and 2011). This property is shown to hold for three candidates or more when May's axioms are invoked together with axioms that rule out the Condorcet and Arrow paradoxes.

The basic *input* is an electorate's opinion profile that may be represented as an m by n matrix of grades when there are m candidates and n voters. A *method of ranking* $\succeq$ is an asymmetric binary relation that compares any two candidates. The *outcome* is a ranking of the candidates. Such methods should meet the following demands.

Axiom 1 (Based on measures) *A voter's opinion is expressed by evaluating each candidate in a scale of grades* Γ*.*

Axiom 2 (Anonymous) *Interchanging the names of voters does not change the outcome.*

Axiom 3 (Neutral) *Interchanging the names of candidates does not change the outcome.*

A method that is both anonymous and neutral is *impartial*.

Axiom 4 (Monotone) *If* $A \succeq B$ *and one or more of* A*'s grades are raised then* $A \succ B$*.*

Axiom 5 (Complete) *For any two candidates either* $A \succeq B$ *or* $A \preceq B$ *(or both, implying* $A \approx B$*).*

Axioms 1 through 5 exactly parallel May's five axioms with one difference, the inputs: for May they are comparisons, here they are measures. With more than two candidates the Condorcet and Arrow paradoxes must be excluded and can be (as was recognized earlier in the context of interpersonal comparisons of preferences [12, 21]).

Axiom 6 (Transitive) *If* $A \succeq B$ *and* $B \succeq C$ *then* $A \succeq C$*.*

Axiom 7 (Independence of irrelevant alternatives (IIA)) *If* $A \succeq B$ *then whatever other candidates are either dropped or adjoined* $A \succeq B$*.*

Theorem 1 *A method of ranking* $\succeq$ *respects domination for three candidates or more if it satisfies Axioms 1 through 7.*

Proof. To compare two competitors among many it suffices to compare them alone by Axiom 7 (IIA). Suppose two candidates A and B have the same set of grades, so that B's list of n grades is a permutation σ of A's list. Compare, first, A with a candidate A',

Opinion profile 1:

	v_1	$\cdots$	$v_{\sigma 1}$	$\cdots$	v_n
$A:$	α_1	$\cdots$	$\alpha_{\sigma 1}$	$\cdots$	α_n
$A':$	$\alpha_{\sigma 1}$	$\cdots$	α_1	$\cdots$	α_n

where A''s list is the same as A's except that the grades given by voters v_1 and $v_{\sigma 1}$ have been interchanged. Suppose $A \succeq A'$. Interchanging the votes of the voters v_1 and $v_{\sigma 1}$ yields the profile

Opinion profile 2:

	$v_{\sigma 1}$	$\cdots$	v_1	$\cdots$	v_n
$A:$	$\alpha_{\sigma 1}$	$\cdots$	α_1	$\cdots$	α_n
$A':$	α_1	$\cdots$	$\alpha_{\sigma 1}$	$\cdots$	α_n

Nothing has changed by Axiom 2 (anonymity), so the first row of Profile 2 ranks at least as high as the second. But by Axiom 3 (neutrality) $A' \succeq A$, implying $A \approx A'$. Thus

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) \approx (\alpha_{\sigma 1}, \alpha_2, \ldots, \alpha_n),$$

and the second list agrees with B's for the first voter of the list.

Compare, now, A with a candidate A'',

Opinion profile 3:

	$v_{\sigma 1}$	v_2	$\cdots$	$v_{\sigma 2}$	$\cdots$	v_n
$A:$	$\alpha_{\sigma 1}$	α_2	$\cdots$	$\alpha_{\sigma 2}$	$\cdots$	α_n
$A'':$	$\alpha_{\sigma 1}$	$\alpha_{\sigma 2}$	$\cdots$	α_2	$\cdots$	α_n

and interchange the votes of voters v_2 and v_{σ_2}. As before and by Axiom 6 (transitivity),

$$(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n) \approx (\alpha_{\sigma 1}, \alpha_{\sigma 2}, \alpha_3, \ldots, \alpha_n),$$

the second list agreeing with B's for the first two voters. Repeating this reasoning shows $(\alpha_1, \alpha_2, \ldots, \alpha_n) \approx (\alpha_{\sigma 1}, \alpha_{\sigma 2}, \ldots, \alpha_{\sigma n})$, so which voter gave which grade has no significance. Therefore the candidates' distributions of grades—called the *merit profile*—is the relevant input for any method that satisfies the seven Axioms, and a candidate's grades has a unique representation when they are listed from the highest to the lowest.

Suppose that A's grades α dominates B's grades β, both given in order of decreasing grades. Domination means that $\alpha_j \succeq \beta_j$ for all j, with at least one strictly above the other. If $\alpha_k \succ \beta_k$ replace β_k in β by α_k to obtain $\beta^1 \succ \beta$ by monotonicity (Axiom 4). Either $\beta^1 = \alpha$ proving that $\alpha \succ \beta$, or $\alpha \succ \beta^1$. In the second case, do as before to obtain $\beta^2 \succ \beta^1$, and either $\beta^2 = \alpha$, or $\alpha \succ \beta^2$. If $\beta^2 = \alpha$ then $\beta \prec \beta^1 \prec \beta^2 = \alpha$ and transitivity implies $\beta \prec \alpha$. Otherwise, repeating the same argument shows that $\alpha \succ \beta$. ∎

The seven Axioms do not characterise a unique method of ranking. Any point-summing method (characterized in [2], chapter 17) satisfies them and so does majority judgment [2, 4]. However, in the one instance where majority rule on two candidates is incontestable, point-summing methods may disagree with it whereas majority judgment always agrees with it.

Polarized electorates

Are there any circumstances in which the majority rule for two candidates cannot be challenged?

One instance immediately leaps to mind: jury decisions. The goal is to arrive at the truth, the correct decision, either the defendant is guilty or is not guilty, there is no gradation of opinion. In this context Condorcet's jury theorem—which says that if each judge has an independent probability of over 1/2 of being correct the majority decision approaches certainty of being correct when the number of judges increases—strongly supports the majority rule. But that context is very different from that of an election between two candidates where gradations of opinion are inherent.

In the context of elections "political polarization" has been given increasing attention in the United States and elsewhere (see e.g., [5, 7]). Polarization means a partisan left/right cleavage in political attitudes that supports ideological extremes and may be attributed to voters, elites, or parties. The concept necessarily concerns an opposition between two candidates or parties. Popular polarization is invoked when large majorities of Democratic and Republican voters are vehemently on opposite sides in their evaluations of issues or candidates. How is this idea to be formalized?

Given an electorate's merit profile consider any two candidates. Many different opinion profiles on the two candidates have this same merit profile (as was seen). The opinion profile on the pair of candidates is polarized when the higher a voter evaluates one candidate the lower the voter evaluates the other. Formally, two candidates A and B are *polarized* when the electorate's opinion profile concerning them satisfies the following condition: if voter v_i evaluates A (respectively, B) higher than does v_j then v_i evaluates B (respectively, A) at most as high as does v_j, for every pair of voters v_i and v_j.

It would seem that it is precisely when an electorate is polarized—or when a jury seeks the correct answer between two opposites—that the "strongly for or strongly against" characteristic of majority voting should render the correct result since there can be no consensus. Consider again the merit profile of Table 2. As was seen, the majority rule elects A with certain opinion profiles and B with others although A's grades dominate B's. The following is the polarized opinion profile having the same merit profile:

	30%	10%	24%	12%	24%
A:	*Very Good*	*Very Good*	*Acceptable*	*Acceptable*	*Poor*
B:	*Poor*	*Acceptable*	*Acceptable*	*Very Good*	*Very Good*

A is the majority rule winner in a 40% to 36% vote, 24% evaluating their merits equally so expressing indifference: for this polarized electorate the majority rule respects the fact that A's grades dominate B's grades.

A method of ranking $\succeq$ is *consistent with the majority rule on polarized pairs of candidates* if both give the identical ranking between every pair of candidates whenever both rankings are decisive. Decisive means no ties.[4]

Point-summing methods are not consistent with the majority rule on polarized pairs. Take for example the two candidates A and B of an electorate that is polarized between them given in Table 4 with the Danish educational system's points, *Very Good* accorded 10 points, *Good* 7 points, and *Poor* 2 points.

	22%	18%	22%	14%	24%
A:	*Very Good*	*Very Good*	*Good*	*Good*	*Poor*
B:	*Poor*	*Good*	*Good*	*Very Good*	*Very Good*

Table 4. Merit profile of a polarized electorate.

A is the MR-winner with 40% of the votes to B's 38%, 22% seeing no difference between them. B is the point-summing winner with $38 \times 10 + 40 \times 7 + 22 \times 2 = 704$ points to A's $40 \times 10 + 36 \times 7 + 24 \times 2 = 700$ points.

Theorem 2 *Any method of ranking $\succeq$ that satisfies Axioms 1 through 7 and is consistent with the majority rule on polarized pairs of candidates must coincide with the majority-gauge rule $\succeq_{MG}$.*

Given an electorate's opinion profile take any pair of candidates. As was seen in the previous proof, the Axioms 1 - 7 imply that only their sets of grades count, i.e., that part of the merit profile that concerns them. The theorem asserts that the majority rule must rank the two candidates of the corresponding polarized opinion profile as does the majority-gauge when both are decisive.

The majority-gauge rule $\succeq_{MG}$ must be explained. A candidate A's *majority-grade* α_A is decided by majority rule: a majority is for at least α_A and a majority is for at most α_A (it is the median of the candidate's grades). Thus, for example, Obama's majority-grade in March 2009 is *Only fair* because a majority of 13+34+30=77% is for at least *Only fair* and 30+23=53% is for at most *Only fair*. A's *majority-gauge* is (p_A, α_A, q_A), where p_A is the percentage of all A's grades strictly above α_A and q_A is the percentage of all A's grades strictly below α_A. Obama's majority-gauge in March 2009 is (47%,*Only fair*,23%). The *majority-gauge rule* is defined by

$$A \succ_{MG} B \text{ when } \begin{cases} \alpha_A \succ \alpha_B \text{ or,} \\ \alpha_A = \alpha_B \text{ and } p_A > \max\{p_B, q_A, q_B\} \text{ or,} \\ \alpha_A = \alpha_B \text{ and } q_B > \max\{p_A, p_B, q_A\}. \end{cases}$$

Here decisiveness means strict majorities for at least and for at most the majority-grades α_A and α_B, and an unequivocal maximum among the p's and q's. With

[4]This restriction eliminates tedious details that are besides the point when there are many voters.

many voters it is almost sure that the majority-gauge rule is decisive (just as it is almost sure that the majority rule is decisive).

The majority-gauge says that A is ranked above B when A's majority-grade is higher than B's. If the majority-grades are equal then among the four blocs of voters who are for a different majority-grade, the largest bloc decides: if it is for a higher grade then that candidate is ranked above the other, if it is for a lower grade then that candidate is ranked below the other.

A suggestive short-cut makes it easy to see the the $\succ_{MG}$ ranking. Adjoin $+p_A$ to a candidate A's majority-grade α_A when $p_A > q_A$ and adjoin $-q_A$ otherwise, then rank them in the natural way: if A's majority-grade is higher than B's then A leads, if both candidates have the same majority-grade then the candidate with the higher adjoined number leads. The majority-gauge shows Obama's performance dropped over the three years:

$$(47\%, \textit{Only fair}+, 23\%) \succ_{MG} (33\%, \textit{Only fair}-, 38\%) \succ_{MG} (33\%, \textit{Only fair}-, 41\%)$$

or

$$(\textit{Only fair} + 47\%) \succ_{MG} (\textit{Only fair} - 38\%) \succ_{MG} (\textit{Only fair} - 41\%)$$

and that A is ranked above B in the example of Table 2 because

$$(40\%, \textit{Acceptable}+, 24\%) \succ_{MG} (36\%, \textit{Acceptable}+, 30\%)$$

or

$$(\textit{Acceptable} + 40\%) \succ_{MG} \textit{Acceptable} + 36\%).$$

Proof. Given an opinion profile with any number of candidates, IIA implies that the order between them must be that determined by the method between two candidates alone. As was shown in the proof of the Theorem 1, the method $\succeq$ depends only on the merit profile, the distributions of the candidates' grades and not on which voters gave them. Thus, the order between them determined by the method will be the same when the electorate's opinion profile is polarized, with A's grades going from highest on the left to lowest on the right and B's from lowest to highest as displayed in Table 5.

$$\begin{array}{c|ccccccccccccc}
 & x_1\% & & & & x_{k-1}\% & & x_k\% & & x_{k+1}\% & & & & x_s\% \\
\hline
A: & \lambda_1^A & \succeq & \cdots & \succeq & \lambda_{k-1}^A & \succeq & \lambda_k^A & \succeq & \lambda_{k+1}^A & \succeq & \cdots & \succeq & \lambda_s^A \\
 & \succ & & \cdots & & \succ & & = & & \prec & & \cdots & & \prec \\
B: & \lambda_1^B & \preceq & \cdots & \preceq & \lambda_{k-1}^B & \preceq & \lambda_k^B & \preceq & \lambda_{k+1}^B & \preceq & \cdots & \preceq & \lambda_s^B \\
\hline
\end{array}$$

A's grades are non increasing and B's non decreasing, so the corresponding grades can be equal at most once (as indicated in the middle line of the profile). Consistency with the majority rule (assumed to be decisive) means one candidate is the MR-winner. Suppose it is A. Then $x_A = \sum_1^{k-1} x_i > \sum_{k+1}^{s} x_i = x_B$.

If $x_A > 50\%$ A's majority-grade is at least λ_{k-1}^A and B's at most λ_{k-1}^B, so A is the MG-winner.

Otherwise, $x_A < 50\%$ (since decisiveness excludes $x_A = 50\%$, though it suffices to assume an odd number of voters). Therefore, $x_A + x_k > 50\%$ and $x_B + x_k > 50\%$, so the candidates' majority-grades are the same $\lambda_k^A = \lambda_K^B =$

λ^*. It must be shown that A's majority-gauge is above B's, or $(p_A, \lambda^*, q_A) \succ (p_B, \lambda^*, q_B)$.

Notice that $p_A \leq x_A$ and $q_B \leq x_A$ but one of the two must be an equality; similarly $q_A \leq x_B$ and $p_B \leq x_B$ but one of the two must be an equality as well. Thus $x_A = \max\{p_A, q_B\}$ and $x_B = \max\{p_B, q_A\} < x_A$. If $p_A = x_A$ then p_A is the largest of the p's and q's, and makes A the MG-winner. If $q_B = x_A$ then q_B is the largest of the p's and q's, and makes B the MG-loser, so A the MG-winner, as was to be shown.

Assume then that A is the MG-winner, $(p_A, \lambda_A, q_A) \succ (p_B, \lambda_B, q_B)$. If $\lambda_A \succ \lambda_B$ the voters who gave the grades λ_A to A and λ_B to B together with all those who gave A at least λ_A (to the left of λ_A in the display) and so gave to B at most λ_B, constitute a majority making A the MR-winner.

So suppose $\lambda_A = \lambda_B$. It is the kth column of the profile $\lambda_k^A = \lambda_k^B$. As before $x_A = \sum_1^{k-1} x_i = \max\{p_A, q_B\}$ and $x_B = \sum_{k+1}^s x_i = \max\{p_B, q_A\}$. Since A is the MG-winner either $p_A > \max\{p_B, q_A, q_B\}$ or $q_B > \max\{p_A, , p_B, q_A\}$. In the first case $x_A = p_A > \max\{p_B, q_A\} = x_B$, in the second $x_A = q_B > \max\{p_B, q_A\} = x_B$. Thus A is the MR-winner. ■

Majority voting between two candidates fails when the winner is strongly rejected by the rest of the electorate whereas the loser is consensual. That opinion profile is excluded in a polarized electorate for there is no consensus. Thus if the majority rule is legitimate in some context it must include polarized electorates. But in polarized electorates the only rules agreeing with the majority rule are those that agree with the majority-gauge, and for all intents and purposes the majority-gauge is decisive when there are many voters.

When the scale of grades is limited to two—*Guilty/Not guilty* or *Approve/Disapprove*—the opinion profile on a pair of candidates is necessarily polarized so approval voting [24, 6] and, of course, majority rule itself agree with the majority-gauge. However, if the scale of grades is more restricted than the set of opinions, making it impossible for a voter to fully express his opinion, the theorem fails. For in this case the majority rule winner against an opponent may not win with the majority rule even when the electorate is polarized between them: with approval voting a voter may *Approve* both candidates or *Disapprove* both without actually being indifferent between them. To guarantee agreement with the majority rule on polarized pairs the scale of grades must faithfully represent the possible diversity of opinion. This encourages a scale of many grades; but for the scale to be practical and held in common by all the voters it should be relatively small. Experimental studies in psychology [18] and experience [2, 4] suggest 7 grades ± 2 are reasonable choices.

Majority judgment is a method of ranking any number of candidates by any number of voters that declares two candidates to be tied only when their sets of grades are identical. However, whenever the majority-gauge is decisive it ranks the candidates exactly as does majority judgment ([2], pp. 236-239). This will almost certainly be the case in an election with many voters, e.g., with over a hundred voters and a scale of six or seven grades (in actual experience to date the majority-gauge has sufficed to determine the rank-order with 19 voters and fewer [4]). Thus for all intents and purposes the majority judgment ranking *is* the majority-gauge ranking when there are many voters.

Measuring intensity

The majority rule fails in the example of Table 2 for a reason that has bothered theorists for centuries [19]: a minority ardently supports A yet 24% of the majority only slightly prefers B to A. Dahl calls this the problem of intensity: "What if the minority prefers its alternative much more passionately than the majority prefers a contrary alternative? Does the majority principle still make sense?" ([9], p. 90). He goes on to say, "If there is any case that might be considered the modern analogue to Madison's implicit concept of tyranny, I suppose it is this one" ([9], p. 99). A more striking example of a potential tyranny of the majority comes from a national poll conducted ten days before the first round of the French presidential election of 2012 by OpinionWay for Terra Nova (see [3]). François Hollande's distribution of grades strongly dominated that of Nicolas Sarkozy (see Table 5).

	Out-standing	*Excel-lent*	*Very Good*	*Good*	*Accept-able*	*Poor*	*To Reject*
Hollande:	12.5%	16.2%	16.4%	11.7%	14.8%	14.2%	14.2%
Sarkozy:	9.6%	12.3%	16.3%	11.0%	11.1%	7.9%	31.8%

Table 5. Merit profile, national poll, French presidential election 2012 [3].

Yet these distributions could have come from the opinion profile of Table 6.

	9.6%	12.3%	11.7%	4.6%	10.2%	5.9%	14.2%
Hollande:	*Exc.*	*V.Good*	*Good*	*Accept.*	*Accept.*	*Poor*	*Rej.*
Sarkozy:	*Outs.*	*Exc.*	*V.Good*	*V.Good*	*Good*	*Accept.*	*Rej.*

	0.8%	5.2%	6.5%	1.4%	5.2%	4.1%	8.3%
Hollande:	*Outs.*	*Outs.*	*Outs.*	*Exc.*	*Exc.*	*V.Good*	*Poor*
Sarkozy:	*Good*	*Accept.*	*Poor*	*Poor*	*Rej.*	*Rej.*	*Rej.*

Table 6. Possible opinion profile, national poll, French presidential election 2012.

Those voters who rate Sarkozy above Hollande (top of profile) do so very slightly, but they represent 54.3% of the electorate whereas Holland is only preferred by 31.5%. This unrealistic opinion profile simply shows how badly majority rule can measure. In the actual run-off Hollande defeated Sarkozy with a bare 51.6% of the votes though it seems his evaluations easily dominated Sarkozy's. This suggests that with both candidates having the same distributions of evaluations—Hollande's dominating Sarkozy's—majority rule might have elected Sarkozy. Another example of how badly majority rule measures is Jacques Chirac's defeat of Jean-Marie Le Pen in the French presidential election of 2002: with but

19.9% of the votes in the first round his 82% in the second round in no way measured his support in the nation at large.

Dahl asks, “If a collective decision is involved, one that requires voting, would it be possible to construct rules so that an apathetic majority only slightly preferring its alternative could not override a minority strongly preferring its alternative?” ([9], p. 92). To attack this question he proposes using an ordinal “intensity scale” obtained “simply by reference to some observable response, such as a statement of one’s feelings ...” ([9], p. 101). He argues that it is meaningful to do so: “I think that the core of meaning is to be found in the assumption that the uniformities we observe in human beings must carry over, in part, to the unobservables like feeling and sensation” ([9], p. 100). This is precisely the role of Axiom 1 and relates to a key problem raised by measurement theorists, the *faithful representation problem*: when measuring some attribute of a class of objects or events how to associate a scale “in such a way that the properties of the attribute are faithfully represented ...”([15], p.1). Practice—in figure skating, wine tasting, diving, gymnastics, assessing pain, etc.—has spontaneously and naturally resolved it. However, some social choice theorists continue to express the opinion that a scale of intensities or grades is inappropriate in elections. As a logical consequence they must also reject the validity of the evaluations of Obama’s performance (Table 1).

Why should intensities be valid—indeed, be necessary—in judging competitions but not in elections? The validity of using intensities as inputs in voting as versus using rankings as inputs is not a matter of opinion or of mathematics, it is at once a philosophical and an experimental issue. Experiments ([2], pp. 9-16 and chapter 15; [11]) have convinced us that nuances in evaluations are as valid for candidates in elections as they are for figure skaters, divers or wines in competitions, though the criteria and scales of measures must be crafted for each individually. Others are also convinced: Terra Nova—“an independent progressive think tank whose goal is to produce and diffuse innovative political solutions in France and Europe”—has included the majority-gauge in its recommendations for reforming the presidential election system of France [22].

Given a scale of grades or intensities Dahl states what he believes a proper rule should accomplish more precisely: “The rules must operate so as to permit a minority veto over the majority only in cases where a relatively apathetic majority would, under pure majority rule, be able to override a relatively more intense minority. That is, the rule must be designed to distinguish the case of ‘severe asymmetrical disagreement’ from the other distributions and permit a minority veto in that case only” ([9], p. 103). “Severe asymmetrical disagreement”—left undefined—presumably means political polarization in the usual sense. Dahl clearly stated the challenge but gave no rule.

The majority-gauge is a rule that provides a solution to Dahl’s intensity problem if “severe asymmetrical disagreement” in an electorate means an electorate that is polarized between two alternatives. When there is severe asymmetrical disagreement most voters would give high grades to their party’s candidate, low grades to the opposing party’s candidate, so that electorate’s opinion profile would be polarized in the mathematical sense. For example, the polarized opinion profile whose grade distributions are those obtained by Holland and Sarkozy (Table 5) above is given in Table 7.

	12.5%	16.2%	3.1%	7.9%	5.4%	5.7%	6.0%
Hollande:	*Outs.*	*Exc.*	*V.Good*	*V.Good*	*V.Good*	*Good*	*Good.*
Sarkozy:	*Rej.*	*Rej.*	*Rej.*	*Poor*	*Accept.*	*Accept.*	*Good*

	5.0%	9.8%	6.5%	7.7%	4.6%	9.6%
Hollande:	*Accept.*	*Accept.*	*Poor*	*Poor*	*Rej.*	*Rej.*
Sarkozy:	*Good*	*V.Good*	*V.Good*	*Exc.*	*Exc.*	*Outs.*

Table 7. Polarized opinion profile for the merit profile of Table 5.

With the polarized opinion profile the majority rule elects Hollande by the score of 50.8% of the votes to Sarkozy's 43.2%. Notice that 68.1% of the voters give one candidate at least *Very Good* and the other at most *Poor*; and 83.3% give one at least *Very Good* and the other at most *Acceptable*. This accords with the usual perception of a polarized electorate. However, when the electorate is not in severe asymmetrical disagreement it is possible for the majority-gauge to disagree with the majority rule (as was shown earlier). The majority-gauge alters the majority rule more subtly than was directly envisioned by Dahl, though this is the logical outcome of using an "intensity scale" as he proposed.

Any rule, however, raises a second key problem of measurement theory, the *meaningfulness problem* [15, 12, 21]: Given a faithful representation, what analyses of sets of measurements are valid or meaningful? When the scale consists of numbers is it justified to sum them or find their averages? The answer is simple: it is valid to calculate averages when the numbers belong to an interval scale, that is, an increase of 1 unit anywhere in the scale has the same meaning (such as temperature, Celsius or Fahrenheit, or length, inches or meters). In a scale (say) of twenty-one numbers going from 0 up to 10 increasing by $\frac{1}{2}$'s, as is used in diving competitions, increasing the score of a competitor from 5 to $5\frac{1}{2}$ is easy whereas increasing it from $9\frac{1}{2}$ to 10 is very hard: an extra point in the middle of the scale is much easier to gain and has much less significance than an extra point at the high end of the scale, so it is not an interval scale. Most methods that rank competing entities use a point-summing method, but virtually all are based on measures that are not interval scales. Denmark's educational system (see above) is an exception: it used extensive data on the distributions of students' past performances to estimate numbers that would result in an interval scale (for an extended discussion of this point (see [2], pp. 171-174, or [4]).

In short, basing the outcomes only on comparisons is at one extreme, asking too little by denying the existence of a scale intensity measures. Taking a cardinal scale of grades as do point-summing methods is at the opposite extreme, it asks too much: in the language of measurement theory it is meaningless to compute sums and averages as do virtually all practical applications (even when the scale is carefully defined), unless the grades belong to an interval scale. Majority judgment and the majority-gauge, based on an ordinal scale, ask for more than comparisons but less than a cardinal scale, so its calculations are meaningful in the sense of measurement theory.

Conclusion

The intent of this article is to make several main points.

- The majority rule for electing one of two candidates is not, with the exception of the very special case of a polarized electorate, a good method.
- Condorcet consistency, in consequence, is not a desirable property, contrary to the widely held view, and should certainly not be considered axiomatic.
- The majority rule says nothing when there is but one candidate; it may fail when there are two; and fails again when there are three or more.
- Comparisons as inputs to methods of voting are insufficient expressions of opinions: they should be replaced by ordinal measures.
- There are methods of ranking that meet May's axioms and Dahl's requirements for any number of candidates—one, two, three or more—when they are based on ordinal measures rather than on comparisons, and they are consistent with the majority rule when it is a good method, namely, for polarized pairs of candidates. When the electorate is large there is only one such method, the majority-gauge.

Methods based on measures or intensities elicit more information from voters, give them the right to express themselves more precisely, and so, it may be hoped, can better determine the will of the electorate.

References

[1] Arrow, Kenneth J. 1951. *Social Choice and Individual Values*. Yale University Press, New Haven, CT. 2nd edition 1963, Wiley, New York, NY.

[2] Balinski, Michel, and Rida Laraki. 2011. *Majority Judgment: Measuring, Ranking, and Electing*. MIT Press, Cambridge, MA.

[3] Balinski, Michel, and Rida Laraki. 2014. "What *should* 'majority decision' mean ?" In J. Elster et S. Novak (eds.), *Majority Decisions*, 103-131. Cambridge University Press.

[4] Balinski, Michel, and Rida Laraki. 2014. "Judge: *Don't vote* !" *Operations Research* 62 483-511.

[5] Barber, Michael, and Nolan McCarty. 2013. "Causes and consequences of polarization." In Jane Mansbridge and Cathie Jo Martin (eds.), *Negotiating Agreements in Politics*, Report of the task force, American Political Science Association, Washington, DC.

[6] Brams, S. J., and P. C. Fishburn. 1983. *Approval Voting*. Birkhauser, Boston, MA.

[7] Brennan Center. 2010. "Ungovernable America? The causes and consequences of polarized democracy." Jorde Symposium. Accessed February 14, 2015, https://www.brennancenter.org/blog/ungovernable-america-jorde-symposium

[8] Condorcet, le Marquis de. 1785. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix.* L'Imprimerie Royale, Paris.

[9] Dahl, Robert A. 1956. *A Preface to Democratic Theory.* The University of Chicago Press, Chicago, IL.

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

[11] Gonzalez-Suitt, J., A. Guyon, T. Hennion, R. Laraki, X. Starkloff, S. Thibault, and B. Favreau. 2014. "Vers un système de vote plus juste?" Département d'Économie, École Polytechnique, Cahier No. 2014-20.

[12] Hammond, Peter J. 1976. "Equity, Arrow's conditions, and Rawls' difference principle." *Econometrica* 44 793-804.

[13] Harris – A Nielson Company. October 2014. "Ratings rebound some for President Obama's job performance." Accessed January 16, 2015, http://www.harrisinteractive.com/NewsRoom/HarrisPolls.aspx

[14] Kemeny, J. G. 1962. "Preference rankings: An axiomatic approach." In J. G. Kemeny and J. L. Snell (eds.), *Mathematical Models in the Social Sciences*, pp. 9-23. Ginn and Co., Boston, MA.

[15] Krantz, David. H., R. Duncan Luce, Patrick Suppes, and Amos Tversky. 1971. *Foundations of Measurement. Vol. 1: Additive and Polynomial Representations.* Academic Press, New York

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

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

[18] Miller, G. A. 1956. "The magical number seven, plus or minus 2: Some limits on our capacity for processing information." *Psychological Review* 63: 81-7.

[19] Morales, Joseph Isidoro. 1797. “Mathematical memoir on the calculation of opinion in elections, 1797.” In Chapter 10, Iain McClean and Arnold B. Urken (eds.), *Classics of Social Choice*, 1995. University of Michigan Press.

[20] Neumann, John von and Oskar Morgenstern. 1944 (3rd edition, 1964). *Theory of Games and Economic Behavior*. Science Editions, New York, NY.

[21] Sen, Amartya. 1977. “On weights and measures: informational constraints in social welfare analysis.” *Econometrica* 45 1539-1572.

[22] Terra Nova. 2011. “Rendre les lections aux lecteurs : le jugement majoritaire.” Accessed February 17, 2015, http://www.tnova.fr/note/rendre-les-lections-aux-lecteurs-le-jugement-majoritaire

[23] Tocqueville, Alexis de. 1831. *De la démocratie en Amérique*, Vol. 1. Edition M.-Th. Gérin, Librairie Médicis, Paris, 1951.

[24] Weber, Robert J. 1977. “Comparison of public choice systems.” Cowles Foundation discussion paper 498. Yale University, New Haven, CT.

[25] Wikipedia. 2014. “Academic grading in Denmark.” Accessed May 7, 2014.

[26] Young, H. P. 1988. “Condorcet’s theory of voting.” *American Political Science Review* 82 1231-1244.

[27] Young, H. P., and A. Levenglick. 1978. “A consistent extension of Condorcet’s election principle.” *SIAM Journal of Applied Mathematics* 35 285-300.