

Série des Documents de Travail

n° 2018-15

**Majority Judgment vs.
Approval Voting**

M.BALINSKI¹

R. LARAKI²

Les documents de travail ne reflètent pas la position du CREST et n'engagent que leurs auteurs.
Working papers do not reflect the position of CREST but only the views of the authors.

¹ CREST; CNRS; Ecole Polytechnique. E-mail : michel.balinski@polytechnique.edu

² CNRS, LAMSADE, Université Paris-Dauphine; PSL; Department of Computer Science, University of Liverpool. E-mail: rida.laraki@dauphine.fr

Majority Judgment vs. Approval Voting

Michel Balinski

CREST and CNRS, École Polytechnique, Palaiseau, France

and

Rida Laraki

CNRS, LAMSADE, Université Paris-Dauphine, PSL, Paris, France

and

Department of Computer Science, University of Liverpool, Liverpool, U.K.

Abstract. Majority judgment (MJ) and approval voting (AV) are compared in theory and practice. Criticisms of MJ and claims that AV is superior are refuted.

The two primary criticisms have been that MJ is not “Condorcet-consistent” and that it admits the “no-show” paradox. That MJ is not Condorcet-consistent is a good property shared with AV: the domination paradox shows majority rule may well err in an election between two. Whereas the no-show paradox is in theory possible with MJ it is as a practical matter impossible. For those who believe this extremely rare phenomenon is important it is proven that MJ with three grades cannot admit the no-show paradox.

In contrast, AV suffers from serious drawbacks because voters can only “tick” or “approve” candidates—at best only *Approve* or *Disapprove* each candidate. With AV voters cannot express their opinions adequately; experiments show that *Approve* is not the opposite of *Disapprove*; and although AV does not admit the no-show paradox it admits the very closely allied “no-show syndrome” and “insensitivity.”

Two is too few. Substantive debate must concern three or more grades.

Key words: Majority judgment, majority rule, approval voting, Condorcet-consistency, domination paradox, no-show paradox, no-show syndrome.

Introduction

In 1925 Walter Lippmann, claimed by many to be the most influential journalist of the twentieth century, pointed to an obvious yet fundamental limitation of the majority rule (MR) as it is practiced in the United States, the United Kingdom, France and most nations throughout the world:

“[W]hat in fact is an election? We call it an expression of the popular will. But is it? We go into a polling booth and mark a cross on a piece of paper for one of two, or perhaps three or four names. Have we expressed our thoughts . . .? Presumably we have a number of thoughts on this and that with many buts and ifs and ors. Surely the cross on a piece of paper does not express them. . . . [C]alling a vote the expression of our mind is an empty fiction.” [25]

Approval Voting (AV)—first proposed in 1977 [41], soon thereafter rashly proclaimed to become “the election reform of the twentieth century” [10]—has been advanced as the method that for a variety of reasons will overcome the limitations of MR [14, 11, 24]. It may be described as a method that asks voters to assign one of two grades, *Approve* or *Disapprove*, to each candidate. The electorate’s rank-order of the candidates is determined by the number of *Approves*. While marking several crosses or ticks rather than the at most one of MR—as AV-voters are typically asked to do—is slightly better than MR, calling an AV-vote an expression of a voter’s mind is an empty fiction as well.

Majority Judgment (MJ) [1], proposed thirty years after AV [41]—a twenty-first century idea—begins by confronting voters with a specific charge, such as (in an election):

“Having taken into account all relevant considerations, I judge, in conscience, that as President of the European Union each of the following candidates would be:”

and then asks them to evaluate each candidate in a scale that contains any fixed, finite number of ordinal grades such as

Outstanding, Excellent, Very Good, Good, Fair, Poor, To Reject.

Majorities determine the electorate’s evaluation of each candidate and the ranking between every pair of candidates—yielding a transitive rank-ordering of all candidates—with the first-placed among them the winner. MJ with more than two grades—six or seven have proven to be good choices—certainly gives voters a much greater ability to express their minds than does AV.

While some AV supporters admit that MJ allows “more nuanced judgments” [11] they steadfastly maintain that AV is superior, advancing a variety of reasons.

The first intent of this article is to present practical and theoretical evidence to show that every one of those “reasons” is either wrong, insignificant, and/or not realized in practice (and some, amusingly, are shared with AV). Among

them—in addition to the lack of Condorcet-consistency [20] and the possibility of the no-show paradox [8, 12, 16, 20]—are that it can elect a candidate preferred by only one voter [11], that it is a “tall order” because too difficult for voters [12], that it induces strategic behavior because “I’m afraid, voters would ... [give] their favorites the maximum grade and their most serious competitors the minimum grade” [12].

AV has been viewed, described, used, or analyzed in three different guises.

(1) It was originally conceived and analyzed in terms of the traditional paradigm of social choice theory, namely, that voters *compare* candidates: implicitly voters rank candidates and draw a line between those who receive a tick or cross and those who do not.¹ With this view changes in the slate of candidatures could induce voters to change their rankings (resulting, for example, from a change in the number of ticks), so with it AV does not escape from Arrow’s impossibility theorem (its advocates made no such claim until it became apparent that AV could be viewed differently).

(2) It has been used as though it were a point-summing method with a tick or cross worth 1 point. Thus, for example, the Social Choice and Welfare Society’s ballot for electing its president had small boxes next to candidates’ names with the instructions: “You can vote for any number of candidates by ticking the appropriate boxes,” the number of ticks determining the candidates’ order of finish. No meaning is ascribed to *Approve* other than it gives 1 point and a candidate’s point total determines the order-of-finish (giving no point is in no way identified with *Disapprove*). It is entirely up to voters to decide how to try to express their opinions.

(3) After MJ became known a new view emerged: “the idea of judging each and every candidate as acceptable or not is fundamentally different” from either voting for a candidate or ranking them ([11], pp.vii-viii). This suggests a belief that voters are able to judge candidates in an ordinal scale of merit with two grades, so AV becomes MJ with a language of two grades (called Approval Judgment in [2]), escapes Arrow’s impossibility theorem, and inherits the other good properties of MJ (except for the limitations due to too few grades). With this point of view MJ vs AV turns into an argument over how many grades MJ should use.

The second intent of the article is to show that *two grades are too few* (however AV is viewed). Two obviously permits no nuance in the expression of voters’ opinions, in and of itself this is sufficiently damning. AV’s behavior also shows that two is too few. It is proven that with MJ the no-show paradox—extremely unlikely in practice with three or more grades—cannot occur when there are three grades, so overcomes all the AV-enthusiasts’ criticisms.

In general, it is argued, majority judgment is a significantly better replacement for majority rule than approval voting has pretended to be.²

¹AV thus viewed is “completely indeterminate” in the following sense: by varying voters’ choices of sincere strategies—moving the cut-off line between *Approve* and *Disapprove* in their preference lists—identical individual preferences can produce every possible order-of-finish so every possible winner. This is true of the election described in section 2.1 [37].

²An August 8, 2018 open message to the membership of INFORMS sent by its President,

1 Methods

A method of ranking candidates (or more generally competitors) based on measures as versus comparisons begins by formulating a question appropriate to its use (in elections to political offices, or in competitions among figure skaters, wines, pianists, or movies) and defining an ordinal scale \succ of merit or evaluation—a *language of grades* Λ —of possible answers. In elections, scales of five to seven ordinal grades have been used; wines have used 21 numerical grades (in Australia) or seven ordinal grades for each of 14 attributes (by the U.I.C.); the Chopin International Piano Competition has used 100 at an elimination stage and 12 for six finalists (for these and other examples see [2]).

Given a language of grades Λ , a method of voting asks voters (or judges of a jury) to evaluate every candidate (or competitor) by assigning each a grade in that scale. The input—called the *opinion profile*³—may be represented as an m by n matrix $\gamma = (\gamma_{ij})$, where m is the number of candidates, n the number of voters, and $\gamma_{ij} \in \Lambda$ the grade assigned to candidate i by voter j .

A *method of ranking* R is a non-symmetric binary relation \succeq_R on candidates that associates to each opinion profile a comparison between them, $A \succeq_R B$ meaning that the electorate prefers candidate A to candidate B or is indifferent between them, $A \approx_R B$ that it is indifferent between them, and $A \succ_R B$ that A is strictly preferred.

1.1 Majority judgment (MJ)

The motivation for MJ was to find a meaningful method of ranking that induces judges and voters to express their evaluations fully and honestly—because an electorate’s or a jury’s collective rank-order of competitors should depend on true opinions and not on strategically calculated expressions—that avoids the bugbears of the traditional theory. It has been described and characterized in several different ways [2, 5, 6]. MJ is based on one single concept that generalizes the idea of median and is not, as has repeatedly been claimed, the median together with “an elaborate set of rules for breaking ties. These are plausible, but there are other tie-breaking rules that would probably work just as well” [12].

Succinctly put,

MJ is majority rule applied to grades,

as the following description shows.

What is the electorate’s *majority* opinion of one candidate with grades cast by n voters? Voters must be treated equally so who gave which grade is of no

Nicholas G. Hall, urged all to vote in the election of INFORMS officers. In it he states that historically only some 20% vote. AV is used although whenever one officer is to be elected there are but two official candidates. Given the arguments of this article the time has perhaps come for INFORMS to reconsider the method it uses to elect its officers.

³The *preference profile* of traditional social choice theory—voters’ rank-orderings of the candidates—may be deduced from the opinion profile when the scale of grades is sufficiently rich.

importance. Define a candidate's *merit profile* to be the set of her grades listed from highest on the left to lowest on the right:

$$\alpha = (\alpha_1, \dots, \alpha_n) \text{ where } \alpha_i \succeq \alpha_{i+1} \text{ for } i = 1, \dots, n-1.$$

For any k less than a majority, $0 \leq k < n/2$, partition the candidate's merit profile into a middlemost block, and left and right blocks of an equal number of grades:

$$\overbrace{\alpha_1 \succeq \dots \alpha_k}^k \succeq \overbrace{\alpha_{k+1} \succeq \dots \alpha_{n-k}}^{n-2k} \succeq \overbrace{\alpha_{n-k+1} \succeq \dots \alpha_n}^k$$

The right and middlemost blocks show there is a majority of $n-k$ for giving the candidate at most the grade α_{k+1} , and the left and middlemost blocks that there is a majority of $n-k$ for giving the candidate at least the grade α_{n-k} : call this a $100(n-k)/n\%$ -majority for $[\alpha_{k+1}, \alpha_{n-k}]$. When $k=0$ this is the unanimous decision for evaluating the candidate at best α_1 and at worse α_n . The larger k the closer to equal are its two grades α_{k+1} and α_{n-k} , so the *more precise* is the majority evaluation. For n odd taking $k=(n-1)/2$ means the middlemost block is reduced to a single grade called the *majority-grade*, the most precise majority evaluation possible.

Consider a real example. The name for a newly formed computer sciences laboratory (the fusion of two existing groups in the spring of 2015) at the Université Paris-Diderot (Paris 7) was to be chosen. Ten names were proposed (here called A through J), 95 persons voted, and they chose a language of five grades⁴:

Good, Rather Good, Not Bad Not Good, Rather Bad, Bad.

A 's merit profile was

	<i>Good</i>	<i>Rather Good</i>	<i>Not Bad Not Good</i>	<i>Rather Bad</i>	<i>Bad</i>
A	35.79%	24.21%	24.21%	10.53%	5.26%

or equivalently, in order to see the middle (designated by \updownarrow),

	<i>Good</i>	<i>Rather Good</i>	\updownarrow	<i>Rather Good</i>	<i>Not Bad Not Good</i>	<i>Rather Bad</i>	<i>Bad</i>
A	35.79%	14.21%	\updownarrow	10.00%	24.21%	10.53%	5.26%

50% of the grades are to the left of the middle, 50% are to the right. There is a 100%-majority for $[Good, Bad]$; and a $50+\epsilon\%$ -majority⁵ for $[Rather Good, Rather Good]$. For any x , $50 < x\% \leq 100\%$, there is a $x\%$ -majority for at most one grade and at least another; e.g., there is a 75%-majority for $[Good, Not Bad Not Good]$. When the two grades are equal they are the competitor's

⁴The actual words were: *Bien, Plutôt bien, Ni bien ni mal, Plutôt mal, Mal.*

⁵Interpret ϵ as one grade.

majority-grade, so A 's majority-grade is *Rather Good*. Naturally the smallest $x\%$ ($> 50\%$) yields the most precise majority decision: with many voters it is essentially certain that they are equal, so the majority-grade.

How does a majority of an electorate rank candidates having sets of grades? Clearly when one candidate's grades are uniformly better than another's (s)he must be ranked above the other.

A candidate A 's merit profile $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ *dominates* B 's merit profile $\beta = (\beta_1, \beta_2, \dots, \beta_n)$ (both written from highest to lowest) when $\alpha_i \succeq \beta_i$ for all i and $\alpha_k \succ \beta_k$ for at least one k ; equivalently, when A has at least as many of the highest grade as B , at least as many of the two highest grades, ..., at least as many of the k highest grades for all k , and at least one "at least" is "more."⁶

Every method of ranking should *respect domination* on every pair of candidates: namely, evaluate one above another when the first's grades dominate the other's.

When one candidate's majority-grade is above another's s(he) must be ranked above the other. When both have the same majority-grade—e.g., A and B in the choice of a new name for the laboratory are both rated *Rather Good*—the most precise majority or the smallest x for which there are $x\%$ -majorities for both that differ decides.

	<i>Good</i>	<i>Rather Good</i>	\updownarrow	<i>Good</i>	<i>Not Bad Not Good</i>	<i>Rather Bad</i>	<i>Bad</i>
A	35.79%	14.21%	\updownarrow	10.00%	24.21%	10.53%	5.26%
B	17.89%	32.11%	\updownarrow	01.57%	20.00%	20.00%	8.42%

Table 1. Merit profiles, laboratory names, Université Paris-Diderot, spring 2015.

Both have $x\%$ -majorities for [*Rather Good*, *Rather Good*] for every x , $50 + \epsilon \leq x \leq 51.57$. The smallest x where they differ are the $51.57 + \epsilon\%$ -majorities: for A it is [*Rather Good*, *Rather Good*] and for B [*Rather Good*, *Not Bad Not Good*]; A 's is better than B 's so A must be ahead of B in the MJ-ranking. These $51.57 + \epsilon\%$ -majorities are determined by the smallest of the four among {14.21%, 10.00%, 32.11%, 1.57%}.

An equivalent description leads to an intuitive rule. Let $p_A\%$ and $q_A\%$ be A 's percentages of grades strictly above and strictly below her majority-grade, and likewise for B . $p_A = 50 - 14.21 = 35.79\%$, $q_A = 50 - 40.00 = 10.00\%$, and $p_B = 17.89\%$, $q_B = 48.43\%$. The $51.57 + \epsilon\%$ -majorities are determined by the biggest of these four, and the rule is this: if the biggest is for a higher grade that candidate leads the other, if it is for a lower grade that candidate lags the other.

A simple diagram pictures the situation in general.

⁶In other contexts domination is called first-order stochastic dominance.

$$\begin{array}{rcc}
A\text{'s merit profile:} & \dots & \dots & \underbrace{[\bar{\alpha} \cdots \downarrow \cdots \underline{\alpha}]} & \dots & \dots \\
& & & \circ = = = \circ & & \\
B\text{'s merit profile:} & \dots & \dots & \underbrace{[\bar{\beta} \cdots \downarrow \cdots \underline{\beta}]} & \dots & \dots
\end{array}$$

Candidate A is ahead of B in the MJ-ranking \succ_{MJ} when: A 's and B 's middle-most grades (bracketed around the center, \downarrow) are all equal except for the left- and right-most (as indicated by \circ), $\bar{\alpha} \neq \bar{\beta}$ or/and $\underline{\alpha} \neq \underline{\beta}$, and either

- $[\bar{\alpha} \cdots \downarrow \cdots \underline{\alpha}]$ dominates $[\bar{\beta} \cdots \downarrow \cdots \underline{\beta}]$, or
- $[\bar{\alpha} \cdots \downarrow \cdots \underline{\alpha}]$ is *more consensual*⁷ than $[\bar{\beta} \cdots \downarrow \cdots \underline{\beta}]$, meaning $\bar{\beta} \succ \bar{\alpha} \succeq \underline{\alpha} \succ \underline{\beta}$.

Thus a succinct complete description of how the majority principle ranks is this:

The MJ-ranking \succeq_{MJ} : For each pair of competitors ignore the maximum equal number of highest and lowest grades of their merit profiles so that when the remaining middlemost grades are compared domination or consensus decides.

This rule—the logical outcome of applying the majority principle to grades—is what many seem to have been groping for. For example, the Fédération International de Natation (FINA) traditionally used a point-summing method for diving, with each of five or seven judges assigning a number grade between 0 and 10 in multiples of $\frac{1}{2}$ (the meanings of each carefully defined [18]) to every dive, the sums of grades determining the rankings. Recently, it changed: when there are five judges the highest and lowest grades are eliminated, when there are seven the two highest and two lowest are eliminated, and then the sums of the remaining grades decide the rankings. Had FINA (or the International Skating Union, the International Gymnastics Federation, or others) gone a bit further—eliminating what must be to distinguish a difference and letting domination or consensus decide rather than sums—they would have used MJ.

It is immediately evident that in practice when there are many voters, (1) every candidate will have a majority-grade, and (2) in comparing candidates consensus is almost surely never invoked since that would require ties in the determination of the maximum number of the rule. Accordingly, assume this is the case and define candidate A 's *majority-gauge* (MG) to be $(p_A\%, \alpha_A, q_A\%)$, and likewise for any candidate, with α_A her majority-grade, $p_A\%$ the percentage of her grades strictly above α_A , $q_A\%$ the percentage of her grades strictly below α_A .

The MG-ranking \succeq_{MG} :

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

⁷In other contexts domination is called first-order stochastic dominance and consensus is called second-order stochastic dominance.

Consistent with the usual interpretation, “+” may be attached to a majority-grade α_C when $p_C > q_C$, and “−” otherwise.

With many voters the MG-ranking is for all practical purposes identical to the MJ-ranking.

	<i>Good</i>	<i>Rather Good</i>	<i>Not Bad Not Good</i>	<i>Rather Bad</i>	<i>Bad</i>
<i>A</i>	35.79%	24.21%	24.21%	10.53%	5.26%
<i>B</i>	17.89%	33.68%	20.00%	20.00%	8.42%
<i>C</i>	31.58%	14.74%	24.21%	15.79%	13.68%
<i>D</i>	13.68%	27.37%	17.89%	22.11%	18.95%
<i>E</i>	18.95%	12.63%	20.00%	21.05%	27.37%
<i>F</i>	15.79%	11.58%	20.00%	24.21%	28.42%
<i>G</i>	15.79%	10.53%	16.84%	25.26%	31.58%
<i>H</i>	10.53%	8.42%	8.42%	16.84%	55.79%
<i>I</i>	7.37%	10.53%	9.47%	15.79%	56.84%
<i>J</i>	4.21%	9.47%	9.47%	18.95%	57.89%

Table 2a. Merit profile, names for newly formed computer sciences laboratory, Université Paris-Diderot.⁸

The full merit profile of the vote at the Université Paris-Diderot is given in Table 2a. Of the 45 pair-by-pair comparisons of grades, 40 are dominations, so all reasonable methods should rank them identically. The MJ-ranking (=the MG-ranking) is given in Table 2b.

	<i>p</i>	α	<i>q</i>
<i>A</i>	35.79%	<i>Rather Good</i> −	40.00%
<i>B</i>	17.89%	<i>Rather Good</i> −	48.42%
<i>C</i>	46.32%	<i>Not Bad Not Good</i> +	29.37%
<i>D</i>	41.05%	<i>Not Bad Not Good</i> −	41.06%
<i>E</i>	31.58%	<i>Not Bad Not Good</i> −	48.42%
<i>F</i>	47.37%	<i>Rather Bad</i> +	28.42%
<i>G</i>	43.16%	<i>Rather Bad</i> +	31.58%
<i>H</i>	44.21%	<i>Bad</i> +	−
<i>I</i>	43.16%	<i>Bad</i> +	−
<i>J</i>	42.11%	<i>Bad</i> +	−

Table 2b. MJ-ranking \succ_{MJ} , names for newly formed computer sciences laboratory, Université Paris-Diderot ($\max\{p, q\}$ in bold).⁹

1.2 Majority rule and the domination paradox

Majority rule (MR)—also called plurality rule and first-past-the-post—is the most used method of voting: voters tick the name of one candidate at most, the

⁸The order of the names is chosen to coincide with the MJ=MG-ranking for added clarity.

⁹The name chosen (*A*) was “Laboratoire d’Informatique de Paris Diderot.”

ranking of the candidates is determined by their numbers of ticks (used, e.g., in Great Britain and the USA); if the top-ranked candidate does not obtain an absolute majority of ticks sometimes (e.g., France) a run-off is held between the two top finishers (call this MR+).

That MR and MR+ measures the relative support of candidates badly was recognized long ago by both Condorcet [15] and Borda [9]. Recent elections in the USA and France prove the point. George W. Bush defeated Albert Gore in 2000 due to the presence of Ralph Nader’s candidacy: Bush’s margin of 537 votes in Florida gave him the state’s 25 Electoral College votes; had Nader not been on the ballot most of his voters would have preferred Gore, making him the victor in Florida and the nation. Jacques Chirac was elected France’s president in 2002 because the candidate who would in all likelihood have defeated him face-to-face, Lionel Jospin, was eliminated in a first-round of 16 candidates, resulting in a second-round that pitted Chirac against a candidate that almost any candidate would have defeated, Jean-Marie Le Pen: the presence of candidates with absolutely no hope of being elected eliminated Jospin. Nicolas Sarkozy won France’s 2007 presidential election for much the same reason: François Bayrou would have defeated any candidate face-to-face—he was the Condorcet-winner—but was eliminated in the first-round. A recent study of the 2017 French legislative elections [32] shows that in 19.2% of the 577 elections the presence of three candidates (in run-offs) denied the election of the Condorcet-winner.

What has not heretofore been appreciated is that MR may go badly wrong in face-to-face encounters, i.e., when there are but *two* candidates.

A method is subject to the *domination paradox* if a candidate *A*’s grades dominate *B*’s but the method ranks *B* above *A*.

MR is subject to the domination paradox. A national French presidential poll of 2012 (described more fully below) proves the point. The leading pretenders

	<i>Outstanding</i>	<i>Excellent</i>	<i>Very Good</i>	<i>Good</i>	<i>Fair</i>	<i>Poor</i>	<i>To Reject</i>
Hollande	12.48%	16.15%	16.42%	11.67%	14.79%	14.25%	14.24%
Sarkozy	9.63%	12.35%	16.28%	10.99%	11.13%	7.87%	31.75%

Table 3a. Merit profile, Hollande-Sarkozy, 2012 French presidential poll.

were F. Hollande and N. Sarkozy. Hollande’s grades dominate Sarkozy’s (Table 3a) yet their merit profile could have come from the opinion profile of Table 3b.

	9.63%	12.35%	11.67%	4.61%	10.18%	11.13%	14.24%
Hollande:	<i>Exc.</i>	<i>V.Good</i>	<i>Good</i>	<i>Fair</i>	<i>Fair</i>	<i>Poor</i>	<i>Rej.</i>
Sarkozy:	<i>Outs.</i>	<i>Exc.</i>	<i>V.Good</i>	<i>V.Good</i>	<i>Good</i>	<i>Fair</i>	<i>Rej.</i>
	0.81%	7.87%	3.80%	6.52%	4.07%	3.12%	
Hollande:	<i>Outs.</i>	<i>Outs.</i>	<i>Outs.</i>	<i>Exc.</i>	<i>V.Good</i>	<i>Poor</i>	
Sarkozy:	<i>Good</i>	<i>Poor</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	

Table 3b. Possible opinion profile, Hollande-Sarkozy (giving the merit profile of table 3a), national poll, 2012 French presidential election.

Using MR a voter would tick the candidate (s)he evaluates higher and abstain or vote blank when both candidates are evaluated *To Reject*. This makes Sarkozy the overwhelming MR-winner with 59.57% of the votes to Hollande’s 26.19%, 14.24% rejecting both.

In fact Hollande defeated Sarkozy in the run-off obtaining 51.6% of the votes, whereas the opinion profile of the national poll showed Hollande winning with 53.9% of the votes. These relatively narrow victories when Hollande’s grades completely dominated those of Sarkozy—and Hollande’s majority-grade was *Good*, Sarkozy’s *Fair*—show that MR can easily go wrong in practice. This is not an isolated event.

Another example is the US presidential election of 2016. The Pew Research Center conducted four in-depth national polls during the course of the election. Among the many questions posed to the participants was the following: “What kind of president do you think each of the following would be—a great, good, average, poor, or terrible president?” Pew Research did not realize that this provides the inputs to a method of election; however, the fact that they used it is testimony to their belief that the question is natural and the answers revelatory of the electorate’s opinion. The last of the four polls was conducted shortly before the election in the period October 20-25 with the result given in Table 4a (except for a 1% change to make a point).

	<i>Great</i>	<i>Good</i>	<i>Average</i>	<i>Poor</i>	<i>Terrible</i>
Hillary Clinton	9%	26%	20%	11%	34%
Donald Trump	9%	18%	16%	11%	46%

Table 4a. Merit profiles, 2016 US presidential poll, Pew Research, October 20-25 [31]. (Clinton’s 9% *Great* was in fact 8% and her 26% *Good* was 27%.)

Clinton’s grades very comfortably dominate Trump’s so MJ (together with any reasonable method) makes Clinton the winner. Moreover, Clinton’s majority-grade is *Average* and Trump’s merely *Poor* (with or without the 1% change). Yet their merit profiles could well have come from the opinion profile of Table 4b. With it MR elects Trump with 49% of the votes to Clinton’s 28%, the remaining 23% of the voters judging both to be *Terrible*.

	9%	18%	11%	11%	
Clinton	<i>Good</i>	<i>Average</i>	<i>Poor</i>	<i>Terrible</i>	
Trump	<i>Great</i>	<i>Good</i>	<i>Average</i>	<i>Poor</i>	
	5%	4%	17%	2%	23%
Clinton	<i>Great</i>	<i>Great</i>	<i>Good</i>	<i>Average</i>	<i>Terrible</i>
Trump	<i>Average</i>	<i>Terrible</i>	<i>Terrible</i>	<i>Terrible</i>	<i>Terrible</i>

Table 4b. Possible opinion profile, Clinton-Trump (giving merit profile of table 4a), Pew Research poll, 2016 U.S. presidential election.

The lesson is clear: MR can easily go wrong even in an election with two candidates. This has far reaching consequences because AV (as practiced), Borda’s,

Condorcet’s, and all methods based on voters’ preferences as versus evaluations become MR when only two candidates compete.

A case in point may well be the 2016 U.S. election. Clinton’s and Trump’s merit profiles were remarkably similar throughout the year—in January, March, August and October—despite the many dramatic ups and downs of the campaign as may be seen in Table 4c.

Clinton:	<i>Great</i>	<i>Good</i>	<i>Average</i>	<i>Poor</i>	<i>Terrible</i>
January	11%	24%	18%	16%	31%
March	11%	22%	20%	16%	31%
August	11%	20%	22%	12%	35%
October	8%	27%	20%	11%	34%

Trump:	<i>Great</i>	<i>Good</i>	<i>Average</i>	<i>Poor</i>	<i>Terrible</i>
January	11%	20%	12%	14%	43%
March	10%	16%	12%	15%	47%
August	9%	18%	15%	12%	46%
October	9%	17%	16%	11%	47%

Table 4c. Merit profiles, 2016 U.S. presidential polls, Pew Research [31].

Clinton’s grades dominated Trump’s in January, March and August, and nearly did in October; her majority-grade was *Average* and his *Poor* all four times. So why did Clinton lose? U.S. voters were in open revolt, determined to show their exasperation with politicians. But how, with MR, could they express this disgust other than by voting for Trump? Had they been able to express their opinions more fully—rating Clinton *Poor* or *Terrible* but Trump as *Poor* or *Terrible* as well—Trump’s very narrow MR victories in Florida, Michigan, Wisconsin and Pennsylvania might well have become Clinton’s giving her 307 Electoral College votes to Trump’s 231.

1.3 Approval voting (AV)

AV is a very slight improvement over MR (when there are more than two candidates): voters tick the name of any number of candidates they wish, the ranking of the candidates is determined by their numbers of ticks. That is its traditional point-summing explanation when used.

If, however, AV voters are specifically asked to evaluate candidates in a scale of merit that contains *Approve* and *Disapprove*—contrary to when it is used with voters asked to tick candidates—AV is MJ with a language of two grades. It has a simple description (already evident from the SCW Society’s instructions): taking

$$\lambda_{Ij} = \begin{cases} 1 & \text{when voter } j \text{ assigns } \textit{Approve} \text{ to candidate } I \text{ and} \\ 0 & \text{when voter } j \text{ assigns } \textit{Disapprove} \text{ to candidate } I, \end{cases}$$

AV ranks candidate A above candidate B when:

$$A \succ_{AV} B \text{ when } \sum_j \lambda_{Aj} > \sum_j \lambda_{Bj}.$$

If *Approve* meant *Good* or better to the Université Paris-Diderot voters the AV-ranking would have been

$$A \succ_{AV} C \succ_{AV} E \succ_{AV} B \succ_{AV} F \approx_{AV} G \succ_{AV} D \succ_{AV} H \succ_{AV} I \succ_{AV} J.$$

That is very different from the MJ-ranking. In this case AV and MJ give the identical winner (*A*'s grades largely dominate any other name's grades). However, AV does not respect domination because *F* and *G* are tied and yet *F*'s grades dominate *G*'s (this happens because two grades are too few). However, AV respects *weak domination*: if one candidate's grades dominate another's the first either leads the second or they are tied.

2 MJ vs. AV: practice

How and why AV fails in practice—in particular, fails in comparison with MJ—is shown via several different uses and experiments.

The inputs to voting are ballots filled out by electors. In a majority vote a ballot poses no question. The voter knows at most one candidate may be given a tick, and the candidates' total ticks decide. Similarly, a typical AV-ballot poses no question but specifies that the voter may give as many ticks as he wishes and the total ticks decide (see Table 5a).

*Vote for one or more candidates by ticking in the appropriate circles.
The candidate with the most votes wins.*

<input type="checkbox"/>	Candidate 1
<input type="checkbox"/>	...
<input type="checkbox"/>	Candidate <i>m</i>

Table 5a. Typical AV-ballot.

Accordingly, ticks are given for completely different evaluations of the candidates and there is no sense that giving no tick to a candidate means he is *Disapproved*. Moreover, instead of eliciting an evaluation of the candidates the instructions suggest comparing them.

*Having considered all relevant information, I believe in conscience,
that as President of the United States of America
each of the following candidates would be:*

	<i>Excellent</i>	<i>Very Good</i>	<i>Good</i>	<i>Acceptable</i>	<i>Poor</i>	<i>Terrible</i>
Candidate 1						
...						
Candidate <i>m</i>						

You must check one single grade for each candidate; no check is taken to mean *Terrible*.

Table 5b. Typical MJ-ballot.

The ballot of Table 5b implicitly forces every voter to evaluate every candidate. There are alternatives to handling “missing grades” (see section 13.4 [2]). Among them one “compensates fairly”: calculate the majority-grade α_C of each candidate C with the grades that are given; then adjoin α_C to C 's grades as many times as necessary to obtain the full complement of grades.

A typical MJ-ballot poses a question and elicits an answer in a scale of grades easily understood by voters or judges (see Table 5b).¹⁰ The question emphasizes the distinct meanings of the answers that are to be given by the voters and that are by and large shared among them. In contrast, the meanings of the ticks or the *Approves* given by a voter in AV-voting vary widely.

2.1 1999 SCW Society 1999

AV was used in the 1999 election of the Social Choice and Welfare Society's president [13, 35]. Members were also asked to indicate their preferences among the three candidates (called A , B , and C). Seventy-one members voted, giving the outcome (where numbers attached to candidates indicate their AV-scores)

$$A(32) \succ_{AV} C(30) \succ_{AV} \succ_{AV} B(14).$$

Two voters ticked three candidates; three ticked two; 64 ticked one; two none.

Fifty-two members gave their preferences

$$\begin{array}{lll} 13 : A \succ B \succ C & 11 : A \succ C \succ B & 9 : B \succ C \succ A \\ 11 : C \succ A \succ B & 8 : C \succ B \succ A & \end{array}$$

Their AV-scores were

$$A(22) \succ_{AV} C(20) \succ_{AV} B(14),$$

giving the same result; here one voter ticked two candidates, 49 ticked one candidate, and 2 ticked none. In this case the face-to-face majority rule votes yield a transitive order (called the *Condorcet-ranking*)

$$C \succ_{MR} A \succ_{MR} B,$$

since (where numbers attached to candidates give their votes)

$$C(28) \succ_{MR} A(24) \quad C(30) \succ_{MR} B(22) \quad A(35) \succ_{MR} B(17);$$

whereas Borda's method¹¹ gives the outcome (where the numbers attached to candidates indicate their Borda-scores)

$$A(59/52) \succ_B C(58/52) \succ_B B(39/52).$$

AV elects the Borda-winner not the Condorcet-winner.

¹⁰

¹¹The *Borda-ranking* is determined by a candidate's *Borda-score*, the average of her votes against every opponent.

- It is claimed that “[t]he AV and [Borda] winners generally coincide with the Condorcet winner”¹² and that AV has “a strong propensity to elect Condorcet candidates” [13]. AV cannot guarantee the election of a Condorcet-winner, as this example shows. These claims are often violated when the result between a pair of candidates is “close” (not surprisingly most methods agree when there are strong winners). Experimental evidence shows that AV-winners are often not Condorcet-winners and often not Borda-winners (see [2], Table 19.1, p. 343).

2.2 French presidential election 2002

In the belief that AV was a reasonable method an experiment was conducted on April 21, 2002, in parallel with the official vote (for more details see [2] pp. 329-333). It was realized in five of Orsay’s twelve voting precincts and the one voting precinct of Gy-les-Nonais (Loiret), where 3,346 voters cast official votes (at most one vote for one candidate) and were then asked to participate in the experiment by casting AV-votes. As in subsequent experiments carried out in France permission was obtained to place separate voting booths and urns on the paths of voters who had already cast their official votes: 2,597 voters (78%) participated.

The AV-ballot stated: “The elector votes by placing crosses [in boxes corresponding to candidates]. He may place crosses for as many candidates as he wishes, but not more than one per candidate. The winner is the candidate with the most crosses.”

	Official vote	AV-score	% of all crosses
Lionel Jospin	19.5% (1 st)	40.5% (1 st)	12.9%
Jacques Chirac	18.9% (2 nd)	36.5% (2 nd)	11.6%
Jean-Marie Le Pen	10.0% (3 rd)	14.6% (11 th)	04.6%
François Bayrou	09.9% (4 th)	33.5% (3 rd)	10.7%
Jean-Pierre Chevènement	08.1% (5 th)	30.3% (4 th)	08.1%
Noël Mamère	07.9% (6 th)	28.9% (5 th)	09.2%
Alain Madelin	05.0% (7 th)	21.3% (6 th)	06.8%
Arlette Laguiller	03.7% (8 th)	15.4% (10 th)	04.9%
Christiane Taubira	03.2% (9 th)	18.9% (7 th)	06.0%
Olivier Besancenot	03.1% (10 th)	17.6% (9 th)	05.6%
Corinne Lepage	02.8% (11 th)	17.9% (8 th)	05.7%
Robert Hue	02.7% (12 th)	11.5% (12 th)	03.6%
Jean Saint-Josse	01.7% (13 th)	07.8% (13 th)	02.5%
Christine Boutin	01.3% (14 th)	07.8% (14 th)	02.5%
Bruno Mégret	01.3% (15 th)	07.7% (15 th)	02.4%
Daniel Gluckstein	00.8% (16 th)	04.3% (16 th)	01.4%

Table 6. Votes and ranks, official and AV, Orsay/Gy-les-Nonais, 1st round French presidential election, April 21, 2002 [2].

¹²For a precise mathematical analysis of the relation between Borda and Condorcet outcomes see [36].

The results are given in Table 6. The three major candidates according to popular opinion and the majority vote all lost relative support, the top-ranked candidate Jospin dropping to 12.9%; whereas all the others (except Chevènement) gained, making the final ranking less rather than more definitive. Jospin’s 40.5% to Chirac’s 36.5% does not give him added legitimacy.

The results do not accurately measure the relative standings of Chirac and Le Pen, the run-off candidates. In these six voting precincts Chirac defeated Le Pen with 89.3% of the votes. There are three possible ways to estimate that outcome using the information revealed by AV. The difficulty is how to “count” a ballot that either gives a cross to both candidates or gives no cross to both. Three estimates were made, in all a cross for one candidate and none for the other gave a 1 to the first. Otherwise, (1) crosses to both and crosses to neither gave both 1/2 (expressing indifference); (2) crosses to both gave both 1/2 (indifference), crosses to neither gave both 0 (says nothing); and (3) crosses to both and crosses to neither gave both 0 (saying nothing). These three interpretations gave the following results, all far from the actual result: (1) Chirac 61%, Le Pen 39%, (2) Chirac 79%, Le Pen 21%, (3) Chirac 80%, Le Pen 20%. This suggests that AV measures badly.

This experiment shows how far an electorate can be from having single-peaked preferences.¹³ If the preferences were single-peaked then the crosses on each ballot would have to be in consecutive boxes (relative to the fixed left-right order). With 16 candidates there are 137 such ballots (16 with 1 cross, 15 with 2 crosses, 14 with 3 crosses, . . . , 1 with 16 crosses), so among the 2,597 ballots at most 137 could be different. In fact, there were many more: 813. However, these AV-ballots were used to deduce how the candidates line-up along a left-right spectrum.¹⁴ This is merely statistical, single-peakedness does not hold, so the existence of a Condorcet-winner is not guaranteed. Nonetheless, Condorcet-cycles may be avoided when sufficiently many voters’ preferences are single-peaked with respect to a fixed order of the candidates (see, e.g., [28, 19, 22]).

- After citing an example where with plurality voting a winning candidate received less than a third of the votes but would have had 54% with AV, it is claimed that it “often will . . . confer legitimacy on their victories to the extent that it shows their support to be widespread” ([14], p. 8). On the contrary, with many candidates spanning a wide and polarized political spectrum AV will often fail to clearly distinguish a winner; moreover, the winner will often fail to be conferred the supposed “legitimacy” of an AV-score close to or above 50%.

- AV measures badly because of its two grades or its limitation to ticks.

This experiment convinced us that AV is a flawed method—completely changing our earlier view—and spurred the search for another method.

¹³An electorate’s preferences are single-peaked if the candidates can be listed in a fixed order from left to right so that *every* voter prefers some one candidate C and the more distant a candidate D is from C to the left or to the right the lower D is in the voter’s preferences.

¹⁴See [2], section 6.2, “Statistical left-right spectra.”

2.3 2009 Louis Lyons Award for Conscience and Integrity in Journalism

As Louis Lyons Award designate, having taken into account all relevant considerations, I believe in conscience that this nominee is:

	<i>Absolutely Outstanding</i>	<i>Outstanding</i>	<i>Excellent</i>	<i>Very Strong</i>	<i>Strong</i>	<i>Commendable</i>	<i>Neutral</i>
Nominee							

Check one grade in the line of each nominee. No check is interpreted to mean *Neutral*.

Table 7a. Ballot for the Louis Lyons Award, 2009 (one line for each nominee).

The Nieman Fellows at Harvard University traditionally decide to whom this award is given. Five highly regarded nominees emerged from a set of many. The Fellows decided to use MJ and the scale of seven grades given in the ballot of Table 7a. The behavior of the judges in this real use of MJ is very similar to that of voters in electoral experiments (see Table 7b): only one judge J_6 gave different grades to all nominees; three judges did not use the highest grade; five gave their highest grade to at least two nominees; and contrary to the predictions of many, judges are far from limiting themselves to the highest or lowest grades.

Judge	C_1	C_2	C_3	C_4	C_5	Judge	C_1	C_2	C_3	C_4	C_5
J_1	4	5	5	3	3	J_2	5	5	6	3	4
J_3	6	5	5	5	5	J_4	6	6	6	6	5
J_5	6	5	4	4	5	J_6	2	5	6	1	3
J_7	6	5	6	5	2	J_8	6	5	5	6	5
J_9	5	0	0	0	0	J_{10}	5	1	5	6	1
J_{11}	5	3	4	3	6	J_{12}	6	6	6	5	5
J_{13}	3	0	0	4	0	J_{14}	6	4	3	0	0
J_{15}	6	5	0	0	0	J_{16}	3	5	6	2	2
J_{17}	6	2	2	2	2	J_{18}	6	5	5	5	5
J_{19}	6	5	4	4	2						

Table 7b. Opinion profile, Louis Lyons Award, 2009.
(6=*Absolutely Outstanding*, 5=*Outstanding*, ..., 0=*Neutral*)

The merit profile is given in Table 7c, and the MJ- and several AV-rankings in Table 7d. Five nominees implies ten comparisons: eight of them are dominations. The two presumably reasonable interpretations for *Approve*—at least *Excellent* or at least *Very Strong*—result in a tie between C_2 and C_3 . The interpretation at least *Outstanding* results in a tie between C_4 and C_5 —whereas C_4 's grades dominate C_5 's—showing that AV does not respect domination.

	<i>Absolutely Outstanding</i>	<i>Outstanding</i>	<i>Excellent</i>	<i>Very Strong</i>	<i>Strong</i>	<i>Commendable</i>	<i>Neutral</i>
C_1	11	4	1	2	1	—	—
C_2	2	11	1	1	1	1	2
C_3	6	5	3	1	1	—	3
C_4	3	4	3	3	2	1	3
C_5	1	6	1	2	4	1	4

Table 7c. Merit profile, Louis Lyons Award, 2009.

	p	MJ-ranking α	q	AV		
				\succ Outstanding	\succ Excellent	\succ Very Strong
C_1	–	<i>Absolutely Outstanding</i>	8	15	16	18
C_2	2	<i>Outstanding</i>	6	13	14	15
C_3	6	<i>Outstanding</i>	8	11	14	15
C_4	7	<i>Excellent</i>	9	7	10	13
C_5	8	<i>Very Strong</i>	9	7	8	10

Table 7d. Rankings¹⁵, Louis Lyons Award, 2009 ($\max\{p, q\}$ in bold)

- With few voters AV often ends in ties although the grades of one competitor dominate another’s. Once again, two grades are simply insufficient.

2.4 Socialist presidential primary, Alfortville 2011.

For the first time ever the French Socialist Party held an open public primary to name their candidate for 2012. It used MR+. École Polytechnique students [21] carried out experiments designed to compare MJ and AV in voting bureaus of Alfortville, a town next to Paris. Participants were asked to vote via MJ- and AV-ballots and to answer several questions.

The MJ-ballot charged the voter, “As the Socialist Party’s nominee in the 2012 presidential election, after having taken into account all considerations, I judge in conscience that this candidate would be:” and offered six possible responses, either *Excellent*, *Very Good*, *Good*, *Acceptable*, *Poor*, or *To Reject*.

The AV-ballot stated “For each of the following candidates to be the Socialist Party’s nominee in the presidential election of 2012, I declare:” and offered two possible responses, either *Approve his/her candidacy* or *Disapprove his/her candidacy*. A specific question was asked—not the usual practice—and it made clear that the alternative to *Approve* is *Disapprove*.

	<i>Excellent</i>	<i>Very Good</i>	<i>Good</i>	<i>Acceptable</i>	<i>Poor</i>	<i>to Reject</i>
Hollande	40.1%	34.5%	12.3%	7.0%	2.5%	3.5%
Aubry	33.1%	36.3%	12.7%	7.4%	3.5%	7.0%
Montebourg	12.0%	27.8%	23.9%	13.7%	11.6%	10.9%
Valls	10.9%	17.6%	26.8%	16.2%	12.0%	16.5%
Royal	12.3%	14.8%	25.7%	18.0%	12.0%	17.3%
Baylet	1.4%	4.6%	14.4%	21.1%	29.6%	28.9%

Table 8a. Merit profile, Alfortville, 2011 [21].

The merit profile is given in Table 8a. 468 persons voted officially, of them 292 (62%) participated in the experiment and 284 votes were valid. The winner is immediately clear because Hollande’s grades dominate all the other candidates’ grades; indeed, each candidate’s grades dominate the next candidate’s from top to bottom except for Royal. The results are given in Table 8b, where “AV-score” are the participants’ AV-votes, “Reported votes” are those they declared having been their official FPP votes (284 voters); “Official FPP” are the

¹⁵Borda and MJ yield the same rankings; the Condorcet-ranking exchanges the places of C_2 and C_3 .

actual first-round results (468 voters); and “AV \succeq *Good*” are the AV results if *Approve* means *Good* or better. The MJ- and AV-rankings are the same except that AV puts Royal in 4th place and Valls in 5th.

MJ-ranking	p	α	q	AV-score	AV \succeq <i>Good</i>	Reported votes	Official FPP
(1) Hollande	40.1%	<i>Very Good+</i>	25.3%	87.3%	87.0%	37.7%	39.7%
(2) Aubry	33.1%	<i>Very Good+</i>	30.6%	85.2%	82.0%	29.2%	28.9%
(3) Montebourg	39.8%	<i>Good+</i>	36.2%	64.1%	63.7%	12.5%	12.3%
(4) Valls	28.5%	<i>Good-</i>	44.7%	53.2%	55.3%	10.0%	08.6%
(5) Royal	27.1%	<i>Good-</i>	47.2%	53.5%	52.8%	10.3%	09.7%
(6) Baylet	41.5%	<i>Poor+</i>	28.9%	25.7%	20.4%	00.4%	00.7%

Table 8b. Results, Alfortville, 2011 [21] ($\max\{p, q\}$ in bold).

The top two AV-scores are both overwhelming, depriving the winner of undisputed legitimacy. Majorities *Approve* all candidates (except Baylet, member of another party) as versus more conflictual elections where no candidate is *Approved* by a majority (e.g., the presidential election of 2002, see above).

Statistically the voters behaved as if *Approve* meant *Good* or better, as may be seen by comparing the AV-score with AV \succeq *Good* (Table 8b). However, the ballots reveal that voters’ thresholds at which they begin to *Approve* varied very widely (see Table 8c), only a third of them interpreting *Approve* to mean *Good* or better.

<i>Excellent</i>	<i>Very Good</i>	<i>Good</i>	<i>Acceptable</i>	<i>Poor</i>	<i>to Reject</i>
24	72	94	54	28	12
8.5%	25.4%	33.1%	19.0%	9.9%	4.2%

Table 8c. Thresholds for *Approve*, Alfortville, 2011 [21].

- Voters attribute very different meanings to *Approve* even when it is clear that its opposite is *Disapprove*, so two grades are too few, putting into question the meaningfulness of an AV-score. There is a striking disparity between conflictual or polarized elections and more consensual ones such as primaries: often in the former no candidate passes the threshold of an AV-score above 50%, in the latter many do. Among the 268 who answered which of MJ and AV permitted them to better express their opinions, not surprisingly MJ was cited by 66.8% to AV’s 33.2%.

2.5 French presidential poll 2012.

Terra Nova, a French think-tank, sponsored a national presidential poll carried out by OpinionWay April 12-16, 2012 (before the first-round of the election held on April 22) to compare MJ with other methods. 993 participants voted with MJ and also according to usual practice—first-past-the-post (FPP) followed by a MR run-off between every pair of the five most likely leaders of the first round. Since the results of FPP varied slightly from the actual national percentages on election day (up to 5%) a set of 737 ballots was found for which those tallies are closely matched, whose merit profile is in Table 9a.

	<i>Outstanding</i>	<i>Excellent</i>	<i>Very Good</i>	<i>Good</i>	<i>Fair</i>	<i>Poor</i>	<i>To Reject</i>
F. Hollande	12.48%	16.15%	16.42%	11.67%	14.79%	14.25%	14.24%
F. Bayrou	2.58%	9.77%	21.71%	25.24%	20.08%	11.94%	08.69%
N. Sarkozy	9.63%	12.35%	16.28%	10.99%	11.13%	7.87%	31.75%
J.-L. Mélenchon	5.43%	9.50%	12.89%	14.65%	17.10%	15.06%	25.37%
N. Dupont-Aignan	0.54%	2.58%	5.97%	11.26%	20.22%	25.51%	33.92%
E. Joly	0.81%	2.99%	6.51%	11.80%	14.65%	24.69%	38.53%
P. Poutou	0.14%	1.36%	4.48%	7.73%	12.48%	28.09%	45.73%
M. Le Pen	5.97%	7.33%	9.50%	9.36%	13.98%	6.24%	47.63%
N. Arthaud	0.00%	1.36%	3.80%	6.51%	13.16%	25.24%	49.93%
J. Cheminade	0.41%	0.81%	2.44%	5.83%	11.67%	26.87%	51.97%

Table 9a. Merit profile,¹⁶ 2012 French presidential poll (737 ballots) [4].

MJ-ranking	p	α	q	$\geq Exc.$	AV-score $\geq V.Good$	$\geq Good$	Official FPP
Hollande	45.05%	<i>Good+</i>	43..28%	(1) 28.63	(1) 45.05	(2) 56.72	(1) 28.6%
Bayrou	34.06%	<i>Good-</i>	40.71%	(5) 12.35	(3) 34.06	(1) 59.30	(5) 09.1%
Sarkozy	49.25%	<i>Fair+</i>	39.62%	(2) 21.98	(2) 38.26	(3) 49.25	(2) 27.3%
Mélenchon	42.47%	<i>Fair+</i>	40.43%	(3) 14.93	(4) 27.82	(4) 42.47	(4) 11.0%
Dupont-A.	40.57%	<i>Poor+</i>	33.92%	(7) 03.12	(7) 09.09	(7) 20.35	(7) 01.5%
Joly	36.77%	<i>Poor-</i>	38.53%	(6) 03.80	(6) 10.31	(6) 22.11	(6) 02.3%
Poutou	26.19%	<i>Poor-</i>	45.73%	(8) 01.50	(8) 05.98	(8) 13.71	(8) 01.2%
Le Pen	46.13%	<i>Poor-</i>	47.63%	(4) 13.30	(5) 22.80	(5) 32.16	(3) 17.9%
Arthaud	24.83%	<i>Poor-</i>	49.93%	(9) 01.36	(9) 05.16	(9) 11.67	(9) 00.7%
Cheminade	48.03%	<i>Reject+</i>	-	(10) 01.22	(10) 03.66	(10) 09.49	(10) 00.4%

Table 9b. MJ-ranking, AV-scores when *Approve* means at least *Good*, at least *Very Good*, and at least *Excellent*, FPP-scores, 2012 French presidential poll (737 ballots) [4].

Table 9b gives the majority-gauges of the candidates, three AV-scores—when *Approve* means at least *Good*, at least *Very Good*, and at least *Excellent*—and the FPP ranking to show the marked differences among them. Table 9c gives the ten MR face-to-face results between the five principal candidates and the deduced Condorcet- and Borda- rankings.

	Hollande	Bayrou	Sarkozy	Mélenchon	Le Pen	Condorcet rank	Borda score
Hollande	-	51.6%	53.9%	68.5%	64.1%	(1)	59.5%
Bayrou	48.4%	-	56.5%	59.4%	70.5%	(2)	58.7%
Sarkozy	46.1%	43.5%	-	50.5%	65.7%	(3)	51.4%
Mélenchon	31.5%	40.6%	49.5%	-	59.7%	(4)	45.3%
Le Pen	35.9%	29.5%	34.3%	40.3%	-	(5)	35.0%

Table 9c. MR results among five candidates, Condorcet- and Borda-rankings, French presidential poll 2012 [4].

The FPP ranking is different from all others, showing how the lack of information concerning the electorate’s evaluations can lead to an outcome that departs widely from the collective consensus. The extreme rightist Marine Le

¹⁶The row sums may differ from 100% due to round-off errors.

Pen is 3rd though she is *Rejected* by 47.63% of the electorate and rated *Poor* or worse by 53.87%, and François Bayrou places 5th although he is much more consensual than either Nicolas Sarkozy, Le Pen or Jean-Luc Mélenchon.

The MJ-, Borda- and Condorcet-rankings are identical among the five major candidates; the three AV-rankings differ. When *Approve* means *Good* or better, Hollande places 2nd (contrary to the usual claim since he is the Condorcet-winner). When *Approve* means *Very Good* or better the winner (Hollande) receives less than a majority of the votes, curtailing his legitimacy. Two meanings of *Approve* place Le Pen in 5th place, one places her 4th, whereas MJ places her 8th: thus AV does better than FPP but does not sufficiently take into account the weight of Le Pen’s negative evaluations. The three meanings of *Approve* place Bayrou 1st, 3rd, and 5th.

- Here three very different methods that use more information than AV agree, whereas AV differs. Two grades are simply not enough.

2.6 French presidential election experiment 2012, Fresne.

AV asks voters to tick those candidates they *Approve*; presumably they *Disapprove* of the others. In the experiment the École Polytechnique students conducted in the context of the Socialist Party’s primary this had been made explicit (see above). If the two grades *Approve* and *Disapprove* constituted a sufficiently rich scale then a voter who does not give a tick to a candidate—who does not *Approve* of the candidate—*Disapproves* of the candidate. Therefore, if instead of asking voters to tick candidates whom they *Approve* they are asked to tick voters whom they *Disapprove* the result should be the same, though here more ticks means lower in the ranking. This proposition was tested.

The experiment was conducted in Fresnes, another town close to Paris, in parallel with the first-round of the French presidential election on April 22, 2012 [21].

In Fresnes’s 12th voting bureau the AV-ballot instructed “Tick the box of each candidate of whom you would APPROVE as President of France,” 771 voted officially, 421 valid MJ- and AV- votes were cast.

Official FPP each bureau	Official FPP		AV		DV	
	Bur. 12	Bur. 14	Bur. #12		Bur. #14	
(1) F. Hollande	39.78%	37.94%	(1)	58.31%	(1)	24.34%
(2) N. Sarkozy	21.35%	19.03%	(4)	25.94%	(4)	50.36%
(3) J-L. Mélenchon	13.67%	15.62%	(2)	42.57%	(3)	36.52%
(4) M. Le Pen	10.60%	12.20%	(6)	12.86%	(10)	75.89%
(5) F. Bayrou	7.83%	9.97%	(3)	33.70%	(2)	28.64%
(6) E. Joly	2.10%	2.61%	(5)	13.53%	(5)	53.94%
(7) N. Dupont-Aignan	1.71%	2.61%	(7)	9.53%	(6)	61.58%
(8) P. Poutou	0.79%	0.77%	(8)	9.09%	(7)	63.72%
(9) N. Arthaud	0.52%	0.46%	(9)	3.99%	(8)	67.54%
(10) J. Cheminade	0.13%	0.31%	(10)	3.77%	(9)	69.69%

Table 10a. Results, official FPP, AV and DisAV (ranks in parentheses), French 2012 presidential election experiment, Fresnes, Bureaus 12 and 14 [21].

In Fresnes’s 14th bureau the DV-ballot instructed “Tick the box of each candidate of whom you would DISAPPROVE as President of France,” 658 voted officially, 408 valid MJ- and DV-votes were cast.

The official FPP together with the AV and DV results of both bureaus are given in Table 10a. The order-of-finish according to FPP is identical in both bureaus (except for a tie between two candidates in Bureau 14); moreover, the actual vote percentages of the candidates in the two bureaus are very similar. By this measure the voting behavior in both bureaus is about the same.

However, the AV and DV results in the two bureaus are different: Le Pen is 6th in one, 10th in the other; Bayrou and Mélenchon swap their 2nd and 3rd places. More to the point, *Approve* and *Disapprove* are not opposites: the candidates’ sums of *Approves* and *Disapproves* are well below 100%, ranging from a low of 62% to a high of 88%, their average under 75%. There are not sufficiently many grades.

	Bureau 12			Bureau 14	
	AV	$AV \succeq \textit{Good}$		DV	$DV \preceq \textit{Poor}$
Hollande	58.31%	55.38%	Hollande	24.34%	19.17%
Mélenchon	42.57%	40.19%	Bayrou	28.64%	24.72%
Bayrou	33.70%	34.40%	Mélenchon	36.52%	30.59%
Sarkozy	25.94%	21.03%	Sarkozy	50.36%	51.22%
Joly	13.53%	14.02%	Joly	53.94%	53.89%
Le Pen	12.86%	10.28%	Dupont-Aignan	61.58%	62.62%
Dupont-Aignan	9.53%	8.18%	Poutou	63.72%	65.05%
Poutou	9.09%	13.55%	Arnaud	67.54%	71.12%
Arnaud	3.99%	2.11%	Cheminade	69.69%	78.64%
Cheminade	3.77%	2.11%	Le Pen	75.89%	71.60%

Table 10b. Actual AV and DV votes compared with $AV \succeq \textit{Good}$ and $DV \preceq \textit{Poor}$, French 2012 presidential election experiment, Fresnes [21].

Statistically the AV results of Bureau 12 are close to interpreting *Approve* to mean *Good* or better and the DV results of Bureau 14 are close to *Disapprove* meaning *Poor* or worse (see Table 10b): *Fair* plays no role whatsoever, again suggesting that two grades do not suffice.

<i>Out-</i>	<i>Excel-</i>	<i>Very</i>		<i>Fair</i>	<i>Poor</i>	<i>To</i>
<i>standing</i>	<i>lent</i>	<i>Good</i>	<i>Good</i>			<i>Reject</i>
Lowest grades of <i>Approved</i> candidates (Bureau 12):						
19.24%	15.91%	30.64%	16.39%	9.26%	2.14%	6.41%
Highest grades of <i>Disapproved</i> candidates (Bureau 14):						
5.45%	5.21%	8.77%	16.11%	22.27%	20.38%	21.80%

Table 10c. Distributions, lowest and highest grades *Approved* and *Disapproved* candidates, French 2012 presidential election experiment, Fresnes [21].

The distributions of the lowest grades given candidates who were *Approved* and the highest grades given candidates who were *Disapproved* are given in Table 10c. As usual voters attribute very different meanings to both *Approve* and *Disapprove*.

The disagreement between AV and DV is easily explained. When a voter is asked to assign *Approve* he assigns them only to those he is “sure” to approve and he evaluates those candidates to whom he gives no ticks at many different levels, a range of opinions that may vary from (say) *Good* down to *To Reject*. Symmetrically, when a voter is asked to assign *Disapproves* he assigns them only to those he is “sure” to disapprove and evaluates candidates to whom he gives no ticks at many different levels, a range of opinions that may vary from (say) *Outstanding* down to *Good*. In either case voters are not able to adequately express their opinions.

The Pew Research Center’s in-depth national polls repeatedly pose the same questions concerning sitting presidents: “Do you approve or disapprove of the way Barack Obama is handling his job as President?” Over Obama’s eight years in office 79 answers are given ([31] p. 78, references are given there to the answers concerning George W. Bush and Bill Clinton). The sums of *Approves* and *Disapproves* are always less than 100%, ranging from a high of 95% to a low of 81% (and similarly for Bush and Clinton).

- *Approve* is not the opposite of *Disapprove*. Two grades are simply too few to adequately express voters’ opinions. At least three grades are necessary.

2.7 Pew Research Center political survey March 2016.

The results of the Pew Research Center’s poll of 1,787 registered voters conducted March 17-27 [30] are given in Table 11a.

	<i>Great</i>	<i>Good</i>	<i>Average</i>	<i>Poor</i>	<i>Terrible</i>	<i>Never heard of</i>
John Kasich	5%	28%	39%	13%	7%	9%
Bernie Sanders	10%	26%	26%	15%	21%	3%
Ted Cruz	7%	22%	31%	17%	19%	4%
Hillary Clinton	11%	22%	20%	16%	30%	1%
Donald Trump	10%	16%	12%	15%	44%	3%

Table 11a. Pew Research Center poll results, March 17-27, 2016 [30].

	<i>p</i>	α	<i>q</i>	$AV \succeq \mathbf{Good}$
(1) John Kasich	33%	<i>Average+</i>	29%	(2) 33
(2) Bernie Sanders	36%	<i>Average-</i>	39%	(1) 36
(3) Ted Cruz	29%	<i>Average-</i>	40%	(4) 29
(4) Hillary Clinton	33%	<i>Average-</i>	47%	(2) 33
(5) Donald Trump	38%	<i>Poor-</i>	47%	(5) 26

Table 11b. MJ- and AV-rankings (*Never heard of* merged with *Terrible*), Pew Research Center poll, March 17-27, 2016 [30].

The MJ- and AV-rankings with *Approve* meaning at least *Good* are given in Table 11b. Sander’s grades very nearly dominate Clinton’s yet AV declares them tied. (When *Approve* means at least *Average* the two rankings are the same.)

- The MJ- and AV-rankings are very different: AV does not sufficiently take into account the fact that 47% of the participants evaluated Clinton as *Poor* or worse (a majority of 62% evaluated Trump as *Poor* or worse). Once again, two grades are too few.

3 MJ vs. AV: theory

What does theory say about the AV-supporters’ criticisms of MJ?

3.1 A winner preferred by only one voter

Brams [12] gives an opinion profile of a jury of five on two candidates to show MJ may elect a candidate preferred by only one voter essentially the same as:

	J_1	J_2	J_3	J_4	J_5
A:	<i>Very Good</i>	<i>Very Good</i>	<i>Very Good</i>	<i>Fair</i>	<i>Fair</i>
B:	<i>Excellent</i>	<i>Excellent</i>	<i>Good</i>	<i>Good</i>	<i>Good</i>

MJ places *A* above *B* because *A*’s majority-grade is *Very Good* whereas *B*’s is *Good*; however, only J_3 prefers *A* to *B*. On the other hand, it is undeniable that a majority believes *A* is *Very Good* and a majority believes *B* is *Good* which happens to be in opposition to the traditional interpretation of majority.

Yet AV can also elect a candidate preferred by only one voter as may be seen with this very same example. AV with *Approve* meaning *Good* or better places *A* over *B* with three ticks to *B*’s two ticks.

3.2 Condorcet-consistency and manipulation

A method that is Condorcet-consistent is necessarily subject to the domination paradox. Moreover, many advocated methods are not Condorcet-consistent when voters express themselves honestly including MJ, AV, the alternative vote, point-scoring methods, and Borda’s method. So why should Condorcet-consistency be considered desirable?

One explanation may be Thomas Paine’s: “[A] long habit of not thinking a thing *wrong*, gives it a superficial appearance of being *right*, and raises at first a formidable outcry in defense of custom. But the tumult soon subsides. Time makes more converts than reason.” [29].

A more serious explanation is the good property that MR with *two* candidates enjoys. It is *strategy-proof* or *incentive compatible*: the best strategy of every voter is to vote honestly. Of course this is a good property of a method only when its results are good with honest behavior (e.g., dictatorship’s strategy-proofness is not an argument in its favor). Honest grades are of very great

importance: it is the true evaluations of the merits of candidates that should be amalgamated, not some exaggerated set of grades whose determination relies on expectations (that in addition, may be erroneous).

So when is MR on two candidates sure to avoid the domination paradox? Roughly said, when the higher a voter evaluates one candidate the lower (s)he evaluates the other. Specifically, a pair of candidates A and B are *polarized* when their opinion profiles have the following property for any two voters v_i and v_j : if v_i evaluates A higher (respectively, lower) than B then v_i evaluates B no higher (respectively, no lower) than v_j . When an electorate is polarized¹⁷ there can be no consensus so the “strongly for or against” characteristic of MR renders the only acceptable result.

	12.48%	16.15%	3.12%	7.87%	5.43%	5.70%	5.97%
Hollande:	<i>Outs.</i>	<i>Exc.</i>	<i>V.Good</i>	<i>V.Good</i>	<i>V.Good</i>	<i>Good</i>	<i>Good</i>
Sarkozy:	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Poor</i>	<i>Fair</i>	<i>Fair</i>	<i>Good</i>
	5.02%	9.77%	6.51%	7.74%	4.61%	9.63%	
Hollande:	<i>Fair</i>	<i>Fair</i>	<i>Poor.</i>	<i>Poor</i>	<i>Rej.</i>	<i>Rej.</i>	
Sarkozy:	<i>Good</i>	<i>V.Good</i>	<i>V.Good</i>	<i>Exc.</i>	<i>Exc.</i>	<i>Outs.</i>	

Table 3c. Polarized opinion profile, Hollande-Sarkozy, 2012 French presidential election poll (for merit profile of Table 3a).

Theorem 1 ([6]) *Majority rule (MR) avoids the domination paradox on pairs of candidates that are polarized.*

The polarized Hollande-Sarkozy opinion profile with the same merit profile as Table 3a is given in Table 3c and the polarized Clinton-Trump opinion profile with the same merit profile as Table 4a is given in Table 4c. The polarized opinion profile makes Hollande the MR-winner with a score of 50.75% to Sarkozy’s 43.28%. Similarly, the polarized profile makes Clinton the MR-winner with 55% to Trump’s 43%. With these opinion profiles MR makes the right decisions.

	9%	26%	11%	9%	2%
Clinton:	<i>Great</i>	<i>Good</i>	<i>Average</i>	<i>Average</i>	<i>Poor</i>
Trump:	<i>Terrible</i>	<i>Terrible</i>	<i>Terrible</i>	<i>Poor</i>	<i>Poor</i>
	9%	7%	18%	9%	
Clinton:	<i>Poor</i>	<i>Terrible</i>	<i>Terrible</i>	<i>Terrible</i>	
Trump:	<i>Average</i>	<i>Average</i>	<i>Good</i>	<i>Great</i>	

Table 4c. Polarized opinion profile, Clinton-Trump, 2016 US presidential election poll, Pew Research Center (for merit profile of Table 4a).

A method of ranking \succeq is *consistent with MR on polarized pairs of candidates* if both give the identical ranking between every such pair.

Theorem 2 ([6]) *A method of ranking that is consistent with MR on polarized pairs (and enjoys several standard properties) must coincide with the MG-rule.*

¹⁷Polarization evokes the single crossing opinion profile of the traditional theory [7, 33].

Since MR is strategy-proof this means the MG-rule shares that property on polarized pairs. It may be shown that MG is the unique method that is strategy-proof (enjoying several standard properties) on this domain. Since MG is MJ when there are many voters this may be said of MJ.

AV, however, is not consistent with MR on polarized pairs. Take, for example, the 2012 French presidential poll (Table 9a). AV with *Approve* meaning *Good* or better elects Le Pen with a score of 32.16% to Poutou's 13.71%. The polarized opinion profile of Le Pen vs. Poutou is given in Table 9d. With it Poutou is the MR-winner with 47.63% of the votes to Le Pen's 46.14%, 6.24% voting for neither (the total of 100.01% is due to round-off error).

	0.14%	1.36%	4.48%	7.73%	12.48%	21.44%	6.24%
Poutou:	<i>Outs.</i>	<i>Exc.</i>	<i>V.Good</i>	<i>Good</i>	<i>Fair</i>	<i>Poor</i>	<i>Poor</i>
Le Pen:	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Poor</i>
	0.41%	13.57%	9.36%	9.50%	7.33%	5.97%	
Poutou:	<i>Poor</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>
Le Pen:	<i>Fair</i>	<i>Fair</i>	<i>Good</i>	<i>V.Good</i>	<i>Exc.</i>	<i>Outs.</i>	

Table 9d. Polarized opinion profile, Le Pen-Poutou, 2012 French presidential election poll (for merit profile of Table 4a).

Nor are point-summing methods consistent with MR on polarized pairs. If *Outstanding* is worth 6 points, *Excellent* 5, . . . , down to *To Reject* 0, Le Pen's 172.75 points easily defeats Poutou's 101.80.

To conclude, MJ is the unique method that is Condorcet-consistent on polarized pairs—where MR works well—and always avoids the domination paradox.

3.3 The no-show paradox and syndrome

Let candidate A 's grades be $\alpha = \{\alpha_1, \dots, \alpha_n\}$ and B 's be $\beta = \{\beta_1, \dots, \beta_n\}$, both written as usual in non-increasing order; and take A^n and B^n to mean those candidates with those n grades.

A method of ranking R is subject to the *no-show paradox* if $A^n \succ_R B^n$, one more voter arrives who evaluates A higher than B , $\alpha_{n+1} \succ \beta_{n+1}$, yet $B^{n+1} \succ_R A^{n+1}$.

Majority judgment admits the no-show paradox (as was pointed out in [2], p. 286) and is easily illustrated. Take the example of section 3.1 but as a merit profile, so that MJ places A above B , and suppose a new voter arrives, rating A *Fair* and B *Poor*, so preferring A to B . This gives the merit profile:

A :	<i>Very Good</i>	<i>Very Good</i>	[<i>Very Good</i>	<i>Fair</i>]	<i>Fair</i>	<i>Fair</i>
B :	<i>Excellent</i>	<i>Excellent</i>	[<i>Good</i>	<i>Good</i>]	<i>Good</i>	<i>Poor</i>

where B 's middlemost block where they differ is more consensual than A 's, so MJ places B above A . This occurs because of a gap in A 's grades at the precise point that causes a reversal, a highly unlikely event in an election with many

voters. However, if the new voter strongly preferred A to B , then the no-show paradox does not occur. For example, if the new voter evaluated A *Excellent* and B *Good*, the merit profile becomes:

A :	<i>Excellent</i>	<i>Very Good</i>	[<i>Very Good</i>	<i>Very Good</i>]	<i>Fair</i>	<i>Fair</i>
B :	<i>Excellent</i>	<i>Excellent</i>	[<i>Good</i>	<i>Good</i>]	<i>Good</i>	<i>Good</i>

A 's middlemost block dominates B 's so MJ places A above B : no no-show paradox.

In general, take a merit profile for which $A^n \succ_{MJ} B^n$ because the extremities of the middlemost blocks where they differ are $[\bar{\alpha}, \underline{\alpha}] \neq [\bar{\beta}, \underline{\beta}]$. Suppose an $(n+1)$ st voter arrives who evaluates A higher than B , $\alpha_{n+1} \succ \beta_{n+1}$.

Theorem 3 *The no-show paradox cannot occur with majority judgment when $\alpha_{n+1} \succeq \max\{\underline{\alpha}, \underline{\beta}\}$ and $\beta_{n+1} \preceq \min\{\bar{\alpha}, \bar{\beta}\}$.*

The no-show paradox cannot occur when (as in the example) the $(n+1)$ st voter sees a real difference between the candidates. It can only occur when that voter either (1) has a relatively good opinion of both candidates ($\alpha_{n+1} \succ \beta_{n+1} \succ \min\{\bar{\alpha}, \bar{\beta}\}$) and B 's evaluation increases, or (2) a relatively bad opinion of both candidates ($\beta_{n+1} \prec \alpha_{n+1} \prec \max\{\underline{\alpha}, \underline{\beta}\}$) and A 's evaluation decreases. However, for the paradox to exist at all it must be assumed that the voter cares more about the victor than about increasing B 's evaluation in the first case or decreasing A 's evaluation in the second case.¹⁸ This assumption is not verified in practice: voters' utilities are unknown and unfathomable (as is further discussed in the conclusions below).

Proof. To begin suppose n is odd and that the first middlemost where A 's and B 's grades differ is the 1st-middlemost. They are singletons, the candidates' majority grades: α_A and β_B with $\alpha_A \succ \beta_B$. Take $\alpha_{n+1} = \alpha_A$ and $\beta_{n+1} = \beta_B$. Then $n+1$ is even and the first middlemost interval where A^{n+1} and B^{n+1} differ is $[\alpha_A, \alpha_A] \neq [\beta_B, \beta_B]$, so $A^{n+1} \succ_{MJ} B^{n+1}$. The monotonicity of majority judgment implies that $A^{n+1} \succ_{MJ} B^{n+1}$ when α_{n+1} is raised or β_{n+1} lowered, so $A^{n+1} \succ_{MJ} B^{n+1}$ for any $\alpha_{n+1} \succeq \alpha_A$ and $\beta_{n+1} \preceq \beta_B$.

So suppose the first middlemost where A 's and B 's grades differ is not a singleton. Recall that in A^n and B^n the middlemost grades strictly within $[\bar{\alpha}, \underline{\alpha}]$ (to the right of $\bar{\alpha}$ and the left of $\underline{\alpha}$) and the corresponding set strictly within $[\bar{\beta}, \underline{\beta}]$ are identical or void. $A^n \succ_{MJ} B^n$ because either (i) $\bar{\alpha} \succeq \bar{\beta}$ and $\underline{\alpha} \succeq \underline{\beta}$ with one comparison strict, or (ii) $\bar{\beta} \succ \bar{\alpha} \succeq \underline{\alpha} \succ \underline{\beta}$.

Case (i). Take any pair $\alpha_{n+1} \succ \beta_{n+1}$ satisfying $\bar{\alpha} \succeq \alpha_{n+1} \succeq \max\{\underline{\alpha}, \underline{\beta}\} = \underline{\alpha}$ and $\underline{\beta} \preceq \beta_{n+1} \preceq \min\{\bar{\alpha}, \bar{\beta}\} = \bar{\beta}$. Such pairs exist, e.g., $\alpha_{n+1} = \bar{\alpha}$ and $\beta_{n+1} = \underline{\beta}$. Place α_{n+1} within the interval $[\bar{\alpha}, \underline{\alpha}]$ and β_{n+1} within the interval $[\bar{\beta}, \underline{\beta}]$ thereby satisfying the property that from left to right the grades are non-increasing. The intervals $[\bar{\alpha}, \underline{\alpha}] \neq [\bar{\beta}, \underline{\beta}]$ remain middlemost intervals since in

¹⁸In the toy-example used to display the paradox the new voter who rates A *Fair* and B *Poor* may be contented having brought down A 's evaluation from [*Very Good*] to [*Very Good, Fair*]: who knows?

A^{n+1} and B^{n+1} the number of grades to their left equals the number of grades to their right. The middlemost grades strictly within $[\underline{\alpha}, \underline{\alpha}]$ (excluding $\underline{\alpha}, \underline{\alpha}$) dominate the corresponding grades strictly within $[\underline{\beta}, \underline{\beta}]$, so $A^{n+1} \succ_{MJ} B^{n+1}$. The monotonicity of majority judgment implies that $A^{n+1} \succ_{MJ} B^{n+1}$ for any $\alpha_{n+1} \succeq \max\{\underline{\alpha}, \underline{\beta}\}$, $\beta_{n+1} \preceq \min\{\bar{\alpha}, \bar{\beta}\}$ and $\alpha_{n+1} \succ \beta_{n+1}$.

Case (ii). Take any pair $\alpha_{n+1} \succeq \beta_{n+1}$ satisfying $\bar{\alpha} \succeq \alpha_{n+1} \succeq \max\{\underline{\alpha}, \underline{\beta}\} = \underline{\alpha}$ and $\underline{\beta} \preceq \beta_{n+1} \preceq \min\{\bar{\alpha}, \bar{\beta}\} = \bar{\alpha}$. Such pairs exist, e.g., $\alpha_{n+1} = \beta_{n+1} = \bar{\alpha}$. Place α_{n+1} within the interval $[\bar{\alpha}, \underline{\alpha}]$ and β_{n+1} within the interval $[\bar{\beta}, \underline{\beta}]$. The intervals $[\bar{\alpha}, \underline{\alpha}] \neq [\bar{\beta}, \underline{\beta}]$ remain middlemost intervals in A^{n+1} and B^{n+1} . If $\alpha_{n+1} \succ \beta_{n+1}$ then, as in case (i), the grades strictly within the interval $[\bar{\alpha}, \underline{\alpha}]$ dominate those strictly within $[\bar{\beta}, \underline{\beta}]$, so $A^{n+1} \succ_{MJ} B^{n+1}$. If $\alpha_{n+1} = \beta_{n+1}$ then $[\bar{\alpha}, \underline{\alpha}] \neq [\bar{\beta}, \underline{\beta}]$ are the first middlemost intervals where the grades of A^{n+1} and B^{n+1} differ, so $A^{n+1} \succ_{MJ} B^{n+1}$. The monotonicity of majority judgment again implies that $A^{n+1} \succ_{MJ} B^{n+1}$ for any $\alpha_{n+1} \succeq \max\{\underline{\alpha}, \underline{\beta}\}$, $\beta_{n+1} \preceq \min\{\bar{\alpha}, \bar{\beta}\}$ and $\alpha_{n+1} \succ \beta_{n+1}$. ■

Theorem 4 *Majority judgment with three grades excludes the no-show paradox.*

Proof. Take the three grades from highest to lowest to be 2, 1, and 0. Suppose (as above) that $A^n \succ_{MJ} B^n$, the first middlemost intervals of A^n and B^n where their grades differ is $[\bar{\alpha}, \underline{\alpha}] \neq [\bar{\beta}, \underline{\beta}]$, and an $(n+1)$ st voter arrives who evaluates A higher than B , $\alpha_{n+1} \succ \beta_{n+1}$.

If n is odd and the first middlemost where A 's and B 's grades differ is the 1st-middlemost, then as was seen they are the candidates' majority grades: α_A and β_B with $\alpha_A \succ \alpha_B$. By Theorem 1 there are only two possibilities where the paradox could occur, namely, when (i) $\beta_{n+1} \prec \alpha_{n+1} \prec \alpha_A$ or (ii) $\alpha_{n+1} \succ \beta_{n+1} \succ \beta_B$.

In the first case three grades implies $\beta_{n+1} = 0, \alpha_{n+1} = 1$, and $\alpha_A = 2$. Consider the worst possible set of evaluations of A and the best of B compatible with the assumptions:

$$\begin{array}{l} A^n : \overbrace{2 \cdots 2}^k \quad \overbrace{\alpha_A = 2}^1 \quad \overbrace{0 \cdots 0}^k \\ B^n : 2 \cdots 2 \quad \beta_B = 1 \quad 1 \cdots 1 \end{array}$$

The $(n+1)$ st voter's grades yields

$$\begin{array}{l} A^{n+1} : \overbrace{2 \cdots 2}^k \quad \overbrace{2 \ 1}^2 \quad \overbrace{0 \cdots 0 \ 0}^k \\ B^{n+1} : 2 \cdots 2 \quad 1 \ 1 \quad 1 \cdots 1 \ 0 \end{array}$$

so $A^{n+1} \succ_{MJ} B^{n+1}$ since A 's two middlemost grades dominates B 's. Thus the no-show paradox does not occur in this most vulnerable of possibilities. The monotonicity of majority judgment then proves that it cannot occur at all.

In the second case $\alpha_{n+1} = 2, \beta_{n+1} = 1$, and $\beta_B = 0$. The worst possible evaluations of A and best of B compatible with the assumptions are:

$$\begin{array}{l} A^n : \overbrace{1 \cdots 1}^k \quad \overbrace{[\alpha_A = 1]}^1 \quad \overbrace{0 \cdots 0}^k \\ B^n : 2 \cdots 2 \quad \overbrace{[\beta_B = 0]}^1 \quad 0 \cdots 0 \end{array}$$

The late voter's grades yields

$$\begin{array}{l} A^{n+1} : \overbrace{2 \ 1 \cdots 1}^k \quad \overbrace{1 \ 1}^2 \quad \overbrace{0 \cdots 0}^k \\ B^{n+1} : 2 \ 2 \cdots 2 \quad 1 \ 0 \quad 0 \cdots 0 \end{array}$$

so again $A^{n+1} \succ_{MJ} B^{n+1}$ and the monotonicity of majority judgment assures that the paradox can never occur.

So suppose the first middlemost where A 's and B 's grades differ is not a singleton. As before, there are but two possibilities for the no-show paradox to arise: either (i) $\beta_{n+1} < \alpha_{n+1} < \max\{\underline{\alpha}, \underline{\beta}\}$ or (ii) $\alpha_{n+1} > \beta_{n+1} > \min\{\bar{\alpha}, \bar{\beta}\}$.

Consider the first case. As was noted before (1) implies $\underline{\alpha} \geq \underline{\beta}$, so with only three grades, $\beta_{n+1} = 0, \alpha_{n+1} = 1$, and $\underline{\alpha} = 2$. Therefore $\bar{\alpha} = 2$, and $\underline{\beta} < 2$ ($\bar{\beta} = 2$ would contradict the hypothesis). The worst possible set of evaluations of A and the best of B compatible with the assumptions are

$$\begin{array}{l} A^n : \overbrace{2 \cdots 2}^k \quad \overbrace{\bar{\alpha} = 2 \cdots 2 \ \underline{\alpha} = 2}^{l \geq 2} \quad \overbrace{0 \cdots 0}^k \\ B^n : 2 \cdots 2 \quad \overbrace{\bar{\beta} = 2 \cdots 2 \ \underline{\beta} = 1}^{l \geq 2} \quad 1 \cdots 1 \end{array}$$

Adjoining the late voter's grades yields

$$\begin{array}{l} A^{n+1} : \overbrace{2 \cdots 2}^k \quad \overbrace{2 \cdots 2 \ 2 \ 1}^{l+1} \quad \overbrace{0 \cdots 0 \ 0}^k \\ B^{n+1} : 2 \cdots 2 \quad 2 \cdots 2 \ 1 \ 1 \quad 1 \cdots 1 \ 0 \end{array}$$

where the $l + 1$ middlemost grades of A dominate the corresponding set of B , showing $A^{n+1} \succ_{MJ} B^{n+1}$. The monotonicity of \succeq_{MJ} proves the no-show paradox cannot occur.

Consider, finally, the second case. With three grades $\alpha_{n+1} > \beta_{n+1} > \min\{\bar{\alpha}, \bar{\beta}\}$ implies $\alpha_{n+1} = 2, \beta_{n+1} = 1$, and $\bar{\beta} = 0$ ($\bar{\alpha}$ cannot be 0 since $A^n \succ_{MJ} B^n$). The worst possible set of evaluations of A and the best of B compatible with the assumptions are

$$\begin{array}{l} A^n : \overbrace{1 \cdots 1}^k \quad \overbrace{\bar{\alpha} = 1 \ 0 \cdots 0 \ \underline{\alpha} = 0}^{l \geq 2} \quad \overbrace{0 \cdots 0}^k \\ B^n : 2 \cdots 2 \quad \overbrace{\bar{\beta} = 0 \ 0 \cdots 0 \ \underline{\beta} = 0}^{l \geq 2} \quad 0 \cdots 0 \end{array}$$

When the $(n + 1)$ st voter adds her votes the grades become

$$\begin{array}{l} A^{n+1} : \overbrace{2 \ 1 \cdots 1}^k \quad \overbrace{1 \ 1 \ 0 \cdots 0}^{l+1} \quad \overbrace{0 \cdots 0}^k \\ B^{n+1} : 2 \ 2 \cdots 2 \quad 1 \ 0 \ 0 \cdots 0 \quad 0 \cdots 0 \end{array}$$

A 's $l + 1$ middlemost grades dominate B 's, so $A^{n+1} \succ_{MJ} B^{n+1}$. The monotonicity of \succeq_{MJ} establishes the result once more. Thus in all cases the no-show paradox cannot occur. ■

There is a more general concept than the no-show paradox that encapsulates the same general idea: not only more support can hurt a candidate but also more support that could be decisive is not decisive.

A method of ranking R is subject to the *no-show syndrome* if $A^n \succeq_R B^n$, one more voter arrives who prefers A to B , $\alpha_{n+1} \succ \beta_{n+1}$, yet $B^{n+1} \succeq_R A^{n+1}$.

AV is subject to the no-show syndrome (as is MJ). Suppose *Approve* means *Very Good* or better with the merit profile

A :	<i>Excellent</i>	<i>Very Good</i>	<i>Good</i>	<i>Fair</i>	<i>Poor</i>
B :	<i>Very Good</i>	<i>Very Good</i>	<i>Fair</i>	<i>Poor</i>	<i>Poor</i>

AV proclaims a tie vote, each candidate receiving two ticks (though A 's grades dominate B 's). One more voter arrives who evaluates A *Good* and B *Fair* or A *Excellent* and B *Very Good*: the vote remains an AV tie. In this case MJ places A ahead of B before and after a sixth voter. Of course this example is in essence the same as that showing MJ admits the no-show paradox: if the late voter sees a real difference between the two candidates—in the MJ case giving A at least *Good* and B at most *Fair* (see Theorem 1), in the AV case giving A a tick and not B —then the no-show syndrome does not occur.

It may be argued that the likelihood for an AV tie in an election with many voters is infinitesimal: true enough, but a late MJ voter is even more rarely in the position of being able to provoke the no-show syndrome or paradox. In the uses and experiments reported on above AV produces ties or near ties whereas MJ does not (unless two candidates have precisely the same set of grades): with few voters AV produces two ties in the Louis Lyons jury; with many voters, AV produces a tie in the Paris-Diderot vote, a near tie in the Alfortville Socialist primary (53.2% for Valls, 53.5% for Royal).¹⁹

AV is subject to another allied property. A method R is *insensitive* if $B^n \succeq_R A^n$, an arbitrarily large set of n' voters arrive who *all* have the same opinion—they give A the grade α and B the grade β with $\alpha \succ \beta$ —yet $B^{n+n'} \succeq_R A^{n+n'}$.

It is evident that AV is insensitive for the simple reason that two grades do not suffice. It occurs when voters see a difference between the two candidates but cannot express it since they either *Approve* both or *Disapprove* both. Insensitivity decreases as the number of grades increases: MJ with sufficiently many grades is sensitive.

AV supporters claim, “By being better able to express their preferences, voters would probably be more likely to go to the polls . . .” ([14] p. 4). There can be no doubt that with a language of six or seven grades voters are able to

¹⁹In experiments comparing methods with 101 ballots AV produced more ties than any of the methods used [2], p. 343.

express their preferences better than with two grades. But more is true. With AV a voter who *Approves* both candidates or *Disapproves* both candidates has no effect on the outcome, so she might as well stay home instead of bothering to vote. With MJ even a voter who gives the same grade to *A* and *B* can (in theory) change the outcome. For example, take the following merit profile where the MJ-winner is *A*:

<i>A</i> :	<i>Very Good</i>	<i>Very Good</i>	<i>Very Good</i>	<i>Fair</i>	<i>Fair</i>
<i>B</i> :	<i>Excellent</i>	<i>Excellent</i>	<i>Good</i>	<i>Good</i>	<i>Good</i>

A voter who gives both *To Reject* makes *B* the MJ-winner. Thus with MJ every voter has the potential to change the outcome. Moreover, many who are opposed to all (or all important) candidates do not vote or cast blank ballots; with MJ they would be incited to participate if only to attribute a *To Reject* to all: were all candidates to have *To Reject* majority-grades there would be little else to do than hold another election with different candidates (a remark that is all the more significant when there is only one candidate).

3.4 Majority judgment with three grades

Majority judgment with only three grades (MJ₃)—say, *Approve*, *Neither (Approve or Disapprove)*, *Disapprove*; or *Competent*, *Harmless*, *Incompetent*—is particularly simple. First, the ballot can be simplified by asking voters to either *Approve* or *Disapprove* or do neither which would mean *Neither* (e.g., Pew Research Center’s presidential approval ratings [30]).

Second, the calculation can be simplified. Let a candidate *A*’s percent (or number) of *Approves* be a_+ and percent (or number) of *Disapproves* be a_- . *A*’s percent of *Neithers* is then $a_0 = (100 - a_+ - a_-)\%$. Define the *score* of a candidate *A* with grades (a_+, a_0, a_-) to be

$$sc(A) = \begin{cases} a_+ & \text{when } a_+ > a_-, \text{ and} \\ -a_- & \text{when } a_+ \leq a_-, \end{cases}$$

in which case the MJ₃ ranking is

$$A \succ_{MJ_3} B \text{ when } \begin{cases} sc(A) > sc(B) \text{ or} \\ sc(A) = sc(B), \text{ and } a_+ - a_- > b_+ - b_-. \end{cases}$$

The second condition says the scores are the same and *A*’s grades dominate *B*’s.

A simple examination of all possibilities proves this is the MJ-ranking.

The Doctoral Council of the Université Paris-Dauphine used MJ₃ in the early summer of 2015. It had one, possibly two, extra fellowships to award and three candidates, one in computer sciences S_c , one in economics S_e , and one in mathematics S_m . Sixteen members J_1, \dots, J_{16} of the Council voted. In keeping with an often used scheme for classifying students in France, the scale of grades used was *A* (highest), *B*, and *C* (lowest). The opinion profile is in Table 12a, the merit profile in Table 12b (the majority-grades are added but are unnecessary to determine the MJ₃-ranking).

	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8
S_e	B	A	A	A	A	A	A	B
S_m	B	C	B	B	A	A	C	A
S_c	A	C	C	A	C	B	B	C

	J_9	J_{10}	J_{11}	J_{12}	J_{13}	J_{14}	J_{15}	J_{16}
S_e	A	B	A	B	A	B	C	C
S_m	C	A	C	C	B	C	A	A
S_c	C	C	B	A	A	A	C	C

Table 12a. Opinion profile, extra fellowship, Doctoral Council, Université Paris-Dauphine, summer 2015.

Grades:	A	B	C	MJ_3 scores	Majority-grade
S_e	9	5	2	+9	A
S_m	6	4	6	-6	B
S_c	5	3	8	-8	C

Table 12b. Merit profile, extra fellowship, Doctoral Council, Université Paris-Dauphine, summer 2015.

S_e dominates S_m which in turn dominates S_c . It is worth noting that this was not an experiment: it was an academic exercise involving the usual stiff competition among disciplines. Nevertheless—despite oft repeated predictions that when the stakes are high voters will use only the highest and lowest grades—voters did not confine themselves to A 's and C 's: 12 B 's were given (together with 20 A 's and 16 C 's). Only five of the 16 voters confined themselves to A 's and C 's. Moreover, although there were three grades and three candidates, only six voters rank-ordered them by using all three grades.

As another example take the 2012 French presidential poll (Table 9a), and suppose that with MJ_3 *Approve* means *Very Good* or better, *Neither* means *Good* or *Fair*, and *Disapprove* means *Poor* or worse, giving the merit profile and MJ_3 scores in Table 13a.

MJ_3 -ranking	<i>Approve</i>	<i>Neither</i>	<i>Disapprove</i>	MJ_3 score	Majority-grade
Hollande	45.05%	26.46%	28.49%	+45.05	<i>Neither</i>
Bayrou	34.06%	45.32%	20.63%	+34.06	<i>Neither</i>
Sarkozy	38.26%	22.12%	39.62%	-39.62	<i>Neither</i>
Mélenchon	27.82%	31.75%	40.43%	-40.43	<i>Neither</i>
Le Pen	22.80%	23.34%	53.87%	-53.87	<i>Disapprove</i>
Dupont-Aignon	9.09%	31.48%	59.43%	-59.43	<i>Disapprove</i>
Joly	10.31%	26.45%	63.22%	-63.22	<i>Disapprove</i>
Poutou	5.98%	20.21%	73.82%	-73.82	<i>Disapprove</i>
Arthaud	5.16%	19.67%	75.17%	-75.17	<i>Disapprove</i>
Cheminade	3.66%	17.50%	78.84%	-78.84	<i>Disapprove</i>

Table 13a. Merit profile three grades, MJ_3 -scores, 2012 French presidential poll (737 ballots).

Five different methods' rankings are given in Table 13b. The rankings all differ. In 34 of the 45 pairs of candidates one's grades dominates the other's. The MJ₃-ranking and all three AV-rankings place Le Pen in 5th place, ignoring the detailed information that places her in 8th place when all the nuances in opinion are taken into account. Three grades in this case do not reflect the fact that of Le Pen's 53.87% *Disapprove*, fully 47.63% are *To Reject*.

	MJ	MJ ₃	AV _≥ <i>Exc.</i>	AV _≥ <i>Very Good</i>	AV _≥ <i>Good</i>
Hollande	1	1	1	1	2
Bayrou	2	2	3	3	1
Sarkozy	3	3	2	2	3
Mélenchon	4	4	4	4	4
Dupont-Aignan	5	6	6	7	7
Joly	6	7	7	6	6
Poutou	7	8	8	8	8
Le Pen	8	5	5	5	5
Arthaud	9	9	9	9	9
Cheminade	10	10	10	10	10

Table 13b. Several rankings, 2012 French presidential poll (737 ballots).

In sum, those AV-enthusiasts who lend a primary importance to the no-show paradox can improve matters by using three grades rather than two, thus giving up AV and using MJ₃: two grades do not suffice. It seems three grades meets their agenda, though three are too few to adequately express voters' opinions in most applications.²⁰

4 Conclusion

Why do voters bother to vote? Few believe that his or her vote in a national contest will change the outcome so “rational” behavior would lead them to abstain. Those voters who go to the polls do so for many reasons: to express their beliefs by giving their opinions (to the extent that the electoral system allows), to participate, to feel they belong to the society in which they live, to be “counted,”

Much of the traditional theory of voting takes a considerably more limited view; namely, that voters care only about who wins an election. It assumes, in the jargon of economic theory, that a voter's utility function depends *only* on the identity of the winner. With this view the extent of the victory, the identity of the runner-up, the place of each candidate in the final standings, and the extent of the support for them, all have nothing to do with voters' wishes.

²⁰Point-summing methods have been characterized as the only methods that avoid the no-show paradox and several other standard properties; however, they should certainly not be used: they are meaningless according to measurement theory and the most manipulable of all methods (see [2]), Chpts. 17 and 19).

In fact nothing is known about voters' wishes and satisfactions—as many elections show—so in a more realistic context “no-show” may be no paradox in theory as well as in practice.

The 2002 AV experiment in Orsay (section 2.2) and polls and experiments with MJ [2, 3] show voters' preferences cannot be explained as single-peaked on a left-to-right political spectrum (although statistically they can be). In theory single-peakedness is very unlikely [22]. A recent study of voting in Switzerland uses principal component analysis to show that at least three dimensions are necessary to map voters' preferences, so the one-dimensional single-peaked condition cannot be met [17].

The 2017 French presidential election shows voters' motivations are far from being limited to the identity of the winner. A first round with 11 candidates, 77.8% of registered voters participating, and 2.6% blank or invalid ballots resulted in four leading candidates: Jean-Luc Mélenchon (far left) 19.58%, Emmanuel Macron (center) 24.01%, François Fillon (right) 20.01%, and Marine Le Pen (extreme right) 21.03%. The second round was between Macron and Le Pen—two very distinct, fundamentally opposed candidates—and yet only 74.6% of the voters turned out, and of all ballots cast fully 11.5% were blank or invalid. Thus 1,536,201 fewer voters participated in the run-off, and of the 35,467,327 who participated 4,085,724 cast blank or invalid ballots: fully 5,622,125 did not wish to chose one or the other candidate in this battle of strongly opposed ideologies, policies, and persons. Many of these voters had voted for Mélenchon or the socialist candidate Benoit Hamon in the first round: they preferred having no voice in the choice of president to giving their vote to a candidate whose ideology was very different than theirs. By preventing them from expressing themselves MR simply disenfranchised them, as would AV as well. The outcome was Macron 66.1%, Le Pen 33.9%: Macron's crushing victory in no way reflected the preference of the electorate for him; instead it expressed the electorate's negative opinion of Le Pen.

A recent study of French parliamentary (1978-2012) and local (2011, 2015) elections also concludes that voters do not care only about who wins. It makes a distinction between “expressive” voters who vote according to their preferences among candidates only and “instrumentally rational” or “strategic” voters who vote based on the likely winners. Their data base leads them to conclude “that a large fraction of voters are . . . ‘expressive’ and vote for their favorite candidate at the cost of causing the defeat of their second-best choice” ([32], p. 42).

In addition to showing what is no more than common sense—that voters must be given the means to express their opinions—this paper has presented evidence—experimental and theoretical—to make a number of key points.

1. MR can easily go astray because voters' opinions cannot be adequately expressed.
2. The two grades or multiple ticks of AV are insufficient expressions of voters' opinions as well.
3. AV's *Approve* is not in practice the complement of *Disapprove* (so two

grades are too few).

4. The no-show paradox—a theoretical possibility with MJ—is insignificant because so unlikely to occur in practice.
5. The no-show paradox is impossible when MJ has three grades.
6. Neither MJ nor AV is Condorcet-consistent when votes are honest, which is a good property of both [6].
7. MJ is consistent with MR on polarized pairs of candidates so resists strategic manipulation on that domain when the grades are sufficiently rich²¹, not true of AV (nor of point-summing methods).
8. In practice voters and judges using MJ do not limit themselves to highest and lowest grades (if they did the results would be those of AV).
9. In practice voters and judges have no difficulty in assigning grades [26]. Experience suggests it is cognitively natural to do so (as versus ranking candidates or partitioning them into two classes). This is shown by repeated experimentation, MJ’s increased use in various instances, including French and American national polls [30], French news sites on the web [34, 38], a grass-roots citizen’s primary conducted in France [23]; and plain common sense.

In short, MJ with six or seven grades (and even with three grades) has striking advantages over AV.

References

- [1] Balinski, Michel, and Rida Laraki. 2007. “A theory of measuring, electing, and ranking.” *Proceedings of the National Academy of Sciences, USA* 104 8720-87225.
- [2] Balinski, Michel, and Rida Laraki. 2011. *Majority Judgment: Measuring, Ranking, and Electing*. MIT Press, Cambridge, MA.
- [3] Balinski, M., 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. Springer, Heidelberg.
- [4] 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, Cambridge, UK.

²¹Sufficiently rich means that a voter who gives the same grade to two candidates considers them of equal merit.

- [5] Balinski, Michel, and Rida Laraki. 2014. “Judge: *Don’t vote !*” *Operations Research* 62 483-511.
- [6] Balinski, Michel, and Rida Laraki. 2016. “Majority judgment vs. majority rule.” Working paper, Laboratoire d’économétrie, École Polytechnique, Palaiseau, France.
- [7] Barberà, S. and B. Moreno. 2011. “Top monotonicity: A common root for single peakedness, single crossing and the median voter result.” *Games and Economic Behavior* 73: 345-359.
- [8] Bogomolny, A. 2011. Review of [2]. Mathematical Association of America. Accessed February 1, 2016.
<http://www.maa.org/press/maa-reviews/majority-judgement-measuring-ranking-and-electing>
- [9] Jean-Charles le Chevalier de Borda. 1784. “Mémoire sur les élections au scrutin.” *Histoire de l’Académie royale des sciences* 657-665.
- [10] Brams, S. J. 1980. “Approval voting in multicandidate elections.” *Policy Studies Journal* 9 102-108.
- [11] Brams, S. J. 2010. “Preface.” In [24], vii-ix.
- [12] Brams, S. J. 2011. “Grading candidates.” (Book review of [2].) *American Scientist* 99 426-427.
- [13] Brams, S. J., and P. C. Fishburn. 2001. “A nail-biting election.” *Social Choice and Welfare* 18 409-414.
- [14] Brams, S. J., and P. C. Fishburn. 1983. *Approval Voting*. Birkhauser, Boston, MA.
- [15] 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.
- [16] Edelman, P. E.. 2012. Book review of [2]. *Public Choice* 151 807-810.
- [17] Etter, E., J. Herzen, M. Grossglauser, and P. Thiran. 2014. “Mining democracy.” In *Proceedings of the Second ACM Conference on Online Social Networks*. ACM, New York, NY, 1-12.
<http://cosn.acm.org/2014/files/cosn060f-etterA.pdf>
- [18] FINA. 2017. “FINA diving rules.” 12-23.
<https://www.fina.org>
- [19] Feld, S. L., and B. Grofman. 1986. “Partial single-peakedness: an extension and clarification.” *Public Choice* 51 71-80.

- [20] Felsenthal, D. S., and M. Machover. 2008. “The majority judgment voting procedure: a critical evaluation.” *Homo oeconomicus* 25 319-334.
- [21] 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.
- [22] Lackner, M.-L., and M. Lackner. 2017. “On the likelihood of single-peaked preferences.” *Social Choice and Welfare* 48 717-745.
- [23] LaPrimaire. 2017).
<https://laprimaire.org>
- [24] Laslier, J.-F., and M. Remzi Sanver (eds.). 2010. *Handbook on Approval Voting*. Springer-Verlag.
- [25] Lippmann, Walter. 1925. “The task of the public.” In *The Phantom Public*, MacMillan Co., New York. Reprinted in *The Essential Lippmann – A Political Philosophy for Liberal Democracy*, C. Rossiter and J. Lare (eds.), Random House, New York, 1963.
- [26] Madhavan, G., C. Phelps, and R. Rappuoli. 2017. “Compare voting systems to improve them.” *Nature* 541 151-153.
- [27] Moulin, H. 1988. “Condorcet’s principle implies the no show paradox.” *Journal of Economic Theory* 45: 53-64.
- [28] Niemi, R. G. 1969. “Majority decision-making with partial unidimensionality.” *American Political Science Review* 63 488-497.
- [29] Paine, Thomas. 1776. *Common Sense*. Project Gutenberg.
<http://www.gutenberg.org/files/147/147-h/147-h.htm>
- [30] Pew Research Center. 2016. “March 2016 political survey.”
<http://www.people-press.org/files/2016/03/03-31-2016-Political-topline-for-release.pdf>
- [31] Pew Research Center. 2016. “As election nears, voters divided over democracy and ‘respect’.”
<http://assets.pewresearch.org/wp-content/uploads/sites/5/2016/11/03170033/10-27-16-October-political-release.pdf>
- [32] Pons, V., and C. Tricaud. 2017. “Expressive voting and its costs: evidence from runoffs with two or three candidates.” Harvard Business School, Working Paper 17-107.
- [33] Puppe, C., and A. Slinko. 2015. “Condorcet domains, median graphs and the single crossing property.” Working paper, July.

- [34] *Rue89*. 2011. “Testez une autre façon de voter à la primaire.” Accessed March 25, 2014. <http://www.rue89.com/making-of/2011/10/05/testez-autre-facon-de-voter-a-la-primaire-225043>
- [35] Saari, D. G. 2001. “Analyzing a nail-biting elections.” *Social Choice and Welfare* 18 415-430.
- [36] Saari, D. G. 2001. *Decisions and Elections: Explaining the Unexpected*. Cambridge University Press, Cambridge.
- [37] Saari, D. G., and J. van Newenhizen. 1988. “Is approval voting an ‘unmitigated evil’?” *Public Choice* 59 133-147.
- [38] Slate.fr. 2011. “Jugement majoritaire: Arnaud Montebourg au second tour.”
<http://www.slate.fr/story/43791/primaire-ps-jugement-majoritaire>
- [39] TNS Sofres. 2007. “Présidentielle 2007: Sondage jour du vote - Tour 1.”
<http://www.tns-sofres.com/etudes-et-points-de-vue/presidentielle-2007-sondage-jour-du-vote-tour-1>
- [40] TNS Sofres. 2012. Accessed March 2, 2016. “Le premier tour de l’élection présidentielle 2012 : sondage jour du vote.”
<https://sites.google.com/a/iepg.fr/trielec/resultats-analyses/enquetes-pre-electorales/vague-t1—jour-du-vote-du-1er-tour/resultatsdusondagetnsssofresavectrielecdu22avril2012>
- [41] Weber, Robert J. 1977. “Comparison of public choice systems.” Cowles Foundation discussion paper 498. Yale University, New Haven, CT.