

**Majority Rule and the Wisdom of Crowds: the Task-Specificity of Majority Rule as
a Predictive Tool**

Hélène LANDEMORE

M.I.T, Masachussets
hlandemore@gmail.com

Prepared for Presentation at the Conference “Majority Decisions,”
Collège de France, Paris, May 13-14, 2009

THIS IS A DRAFT, PLEASE DO NOT CIRCULATE

“Majority Rule and the Wisdom of Crowds: the Task-Specificity of Majority Rule as a Predictive Tool”

Hélène Landemore

Imagine that you are lost in a forest with a group of friends. After some discussion, the group agrees that they should keep walking (as a group) in a given direction but disagrees between two options: going North or going North West. The majority vote to go North. What is the rationale for every single member of the group to go with the majority decision? What is, in the first place, the rationale for using majority rule to reach a collective decision, rather than let the most senior person in the group decide or the one with the most hiking experience? One could argue on the basis of procedural fairness that by so doing the group respects the equality of each individual. As David Estlund has proposed, however, if fairness is the main justification for majority rule, why not flip a coin? Another fair option would be to choose randomly a person in the group and let that person make the decision for the group.¹

Another argument in favor of majority rule in that scenario would be that it maximizes individuals' autonomy. Everybody is given a chance at some amount of self-rule through their vote. If the group ends up going in the wrong direction, at least no member will be able to blame it on anybody but themselves (including the members of the minority, since we assume that the principle of majority rule was embraced by them as well as any result that using that rule might produce).

While these reasons have some force, I would argue that the main reason to embrace majority rule as the collective decision rule in that scenario is that, all other

¹ That would be fair because everybody would have the same probability of influencing the outcome.

things equal otherwise, the use of majority rule maximizes the chances that the group picks the right answer. And the reason to embrace the particular decision produced by majority rule in that case is that it is more likely to be accurate than the prediction of the minority.

In my earlier work (Landemore 2007 and 2009), I have defended the broader claim that, in the political sphere, majority rule (and democracy more generally) is at least partly valuable because of its tendency to produce smart collective decisions under the normal circumstances of politics. In so doing I embraced an “epistemic” argument for majority rule, which can be traced back to the Condorcet Jury Theorem (1785) all the way to contemporary democratic theorists (e.g., Dahl 1982, Cohen 1986, Estlund 1997 and 2009, Goodin 2003). I also connected this argument to a larger literature on “the Wisdom of Crowds” (e.g., Surowiecki 2004, Sunstein 2006, Page 2007). While there exist several competing ways to account for the epistemic properties of majority rule, the general claim I made is that under certain (reasonable) conditions, majority rule maximizes the group’s chance to figure out the right policy platform or candidate or more generally outcome out of two alternatives.

In this paper, I would like to argue more specifically for the task-specificity of majority rule as a prediction tool, by contrast with the task-specificity of deliberation as a problem-solving tool. Whereas there may be cases in which the epistemic properties of the two decision-procedures may genuinely be in competition, I argue that deliberation and majority rule are best seen as sequential and complementary mechanisms of collective decision-making. Both procedures have epistemic properties but they are not best suited to the same tasks. In passing I consider if the view of majority rule as a

predictive tool is compatible with the view of majority rule as a way to adjudicate peacefully and fairly between competing interests and preferences. I conclude that, in practice, the two tasks are often mixed, which might explain why majority rule is a much blunter decision-making tool than it need be.

In the first section, I review the epistemic case for majority rule, pointing out its singular status in the literature on democracy as the most recent and the least used argument for majority rule. I blame this underdog status of the epistemic case for democracy on a general elitist prejudice combined with the apparent shortcomings of the most widespread account of majority rule's epistemic properties, namely the Condorcet Jury Theorem.

In the second section, I turn to competing accounts of the epistemic properties of majority rule and find Page's argument from cognitive diversity particularly compelling.

The third section briefly compare deliberation and majority rule to conclude that while those decision mechanisms may be viewed as competing, they are best conceptualized as sequential and complementary tools of collective decision-making.

The fourth section deals with a couple of objections to this defense of majority rule as a predictive tool. The first one points out the fact that it is hard to see politics as a prediction game when there is no verification mechanism. The second objection argues that politics is about arbitrating between conflicting preferences and values, not making predictions.

1. The Epistemic Case for Majority Rule: the Underdog of Democratic Theory

Majority rule is a decision rule in which more than half the people have the final word on a choice between two or more options. In this paper I focus on the simplified case of two options.² While majority rule is by definition more democratic than minority rule, a decision rule can be used by a small number of people (more than two) or a large number of people. It is democratic only with respect to the group whose members it includes. I make this relatively trivial point to clarify that I am not especially interested here in defending democracy per se, but simply the principle of majority rule, whether it is used by an oligarchy, a group of representatives, or any other group.

Majority rule is celebrated for many reasons, although none of them is fully satisfying. As a commentator could point out a few years ago, “although majority rule finds ready acceptance whenever groups make decisions, there are surprisingly few philosophically interesting arguments in support of it” (Risse 2004:1).³ Existing arguments for majority rule roughly fall into two categories, intrinsic and instrumental. On the intrinsic side, some defend majority rule as a procedure that embodies the principles of political equality, fairness, justice, respect, or freedom. Those who defend

² This might seem like dodging a crucial difficulty since the main problems for majority rule arise where there are three options or more to choose from, as with Arrow’s Impossibility Theorem. This worry is justified, however, only if one accepts the assumptions on which Arrow’s Impossibility Theorem depends, in particular the assumption of unlimited domain of preferences. I refer to Mackie and List for a convincing critique of this assumption (Mackie 2003, List 2001). While I do not deny that extending the number of options beyond three might create problems for the epistemic case presented here, I think it is important to focus first on establishing its plausibility for binary predictions. If the idea that a choice is always binary seems to far-fetched, consider that other political mechanisms, such as some form of pre-voting deliberation, may plausibly generate such a partitioning of political problems into binary options.

³ Risse goes on to list six standard arguments, all of which he finds wanting in some respect. Those arguments include the argument that short of unanimity, the majority must logically prevail; the argument that majority rule maximizes the number of people who exercise self-determination, the argument that majority rule expresses respect for people, especially the minority, in allowing everyone one a say in the final decision; the Condorcet Jury Theorem, May’s Theorem and the argument that the result of majority rule represents a necessary compromise. Political scientists and economists have thus tried to demonstrate a causal connection between democracy and what is empirically referred to as “the democratic advantage”—i.e., the observed superiority of democratic regimes in terms of life-expectancy of their population and economic growth of the corresponding countries (e.g., REF). Such empirical arguments can no doubt support a strong instrumental case for majority rule, but only if and when a causal relation between the use of majority rule and such empirical outcomes can be established.

majority rule on instrumental grounds emphasize that it minimizes conflict and fosters peaceful resolution of disagreement.⁴ A slightly distinct instrumental argument is that majority rule by definition entails minimal harm to individual rights since even in the case of a tyrannical majority, the number of people oppressed is less than half of the group, whereas, by construction, the tyranny of a minority oppresses more than half of the group. There is also an empirical instrumental argument for majority rule, which is conceptually distinct from the previously listed arguments. The empirical instrumental argument connects the use of majority rule and certain valued observable outcomes.⁵

Finally, the last type of instrumental argument, defended in this paper, proposes that majority rule maximizes the group's chances to pick the right answer to a given question out of two options, under a given set of conditions. The "epistemic" argument for majority rule, as it is also known (see Cohen 1989 and Coleman and Ferejohn REF for the initial christening) is instrumental because it is outcome-oriented, rather than procedure-oriented.

I would like to argue that the epistemic argument for democracy is unjustly underrated and definitely underused. Before I start defending it properly, let me venture a few conjectures as to why this is the case. The underdog status of the epistemic argument for majority rule may have to do with the lateness of its formulation, particularly compared to other arguments for majority rule such as those based on the value of equality or liberty. The epistemic argument for majority rule is at most as old as the

⁴ The instrumental arguments sometimes come with empirical backup that existing democracies which rely on majority rule for important decisions such as the election of leaders do well on many quantifiable fronts (such as life expectancy, standards of living etc. Of course it is not absolutely clear whether all those good outcomes are to be connected to the fact that majority rule is resorted to or to other factors.

⁵ It is admittedly methodologically complicated to connect specific empirical outcomes to the workings of a given institution in a democracy.

Condorcet Jury Theorem, which was formulated in 1785. The epistemic case for democracy in general is much older (at least as old as the Sophists) but that is probably because the focus was on the practice of deliberation, not majority rule.⁶ In order to understand that majority rule has epistemic properties, indeed, one needs a mathematical notion of probability which was not available before mid 17th century, when Pascal and Fermat started theorizing about the logic of chance games. The Condorcet Jury Theorem itself was only formulated a century later and its interest was not even properly understood until recently, thanks to its rediscovery by Duncan Black in the 1950s (at about the same time as Arrow formulated his impossibility theorem).⁷

Perhaps because of his recent pedigree, the epistemic case for majority rule occupies a strange position in the contemporary ranking of arguments for majority rule. On the one hand, democratic theorists are happy to use it as an additional argument to justify majority rule used by the legislation (e.g., Waldron 1999). On the other hand, when the group in question is the people at large, that is millions of democratic citizens involved in a referendum for example, the epistemic case for majority rule is hardly ever made. Jeremy Waldron typically prefers to turn to May's theorem,⁸ a theorem emphasizing not the epistemic reliability of the procedure but its procedural fairness (Waldron 1999: 137). More generally, one would be hard pressed to find a single

⁶ In the historical chapter of my doctoral work, I have identified in Spinoza and other earlier authors proto-versions of the Condorcetian argument but none of them could quite point out the reason why majority rule could be relied on to produce good outcomes for lack of the relevant mathematical conceptual tools.

⁷ In all fairness, it is likely that Condorcet's theorem had a huge influence on Madison's Federalist 10 but this is a point only made recently (Schoffield, Unpublished) and which remains largely unacknowledged in contemporary democratic literature. Furthermore, Federalist 10 itself did not seem to spur a great epistemic confidence in the selection of leaders through majority rule, as it was interpreted merely as an argument in favor of dissolving factions in a large pool of applicants.

⁸ According to May's theorem (1952), simple majority rule is the only anonymous, neutral and monotone choice function when there are two candidates. Notice that, like the CJT, May's theorem applies only for two options.

democratic theorist willing to get caught using the epistemic case in support of direct forms of democracy.

One of the reasons for that widespread skepticism about the epistemic case for majority rule may have to do with the fact that until recently the only account of the epistemic properties of majority rule was the CJT, which seems overly sanguine about the performance of large groups. As is well known, the CJT demonstrates that among large electorates voting on some yes or no question, majoritarian outcomes are virtually certain to track the “truth,” as long as three conditions hold: 1) voters are better than random at choosing true propositions (what I call the Enlightenment assumption)⁹; 2) they vote independently of each other (the Independence assumption); and 3) they vote sincerely or truthfully (the Sincere Voting assumption). To briefly illustrate, consider 10 voters, each of which has a .51 probability to be correct on any yes or no question. A majority of 6 will have 52% chance of being right. If you now expand the group to 1000 people and a majority of 600 hundred is almost 100% sure to be right.

According to the theorem, there is thus a simple trade-off between numbers and individual competence.¹⁰ By default large groups have thus an advantage over smaller ones because adding individuals to the group is a “cheaper” way to raise the probability that the group gets to the right answer than trying to increase the political IQ of individuals (their probability of being right on any yes or no question) in the group. In

⁹ Technically the CJT requires that each citizen be better than random at choosing between two options. This condition can however be generalized into the empirically more plausible assumption that the *average* voter has such a competence (see Grofman and Feld 1982).

¹⁰ In this example, one can see that simply expanding the size of the voting group almost certainly guarantees accuracy. In order to obtain the same level of accuracy without increasing the size of the group, one would have to considerably increase the individual probability of being right.

other words, if the choice is between quantity and quality, it seems that the easier way to go is with higher numbers.

Such a conclusion, however, seems hard to accept for a lot of commentators, who criticize the implausibility of the assumptions on which the truth of the theorem depends. Duncan Black set the mood for this reception of the theorem in that respect, since he rejected the very premise of an epistemic argument by worrying that the idea of a “probability of the correctness of a voter’s opinion” was a phrase “without definite meaning” (Black 1958: 163). But even when they admit that this idea makes sense, other commentators find either the Enlightenment assumption, the Independence assumption or the Sincere Voting assumption utterly implausible (e.g., Rawls 1971: 314-15, Lhada 1992; Christiano 1996: 35).¹¹ Even David Estlund, an unambiguous advocate of an (at least partially) epistemic theory of democracy, who repeatedly mentions the CJT in support of the superiority of majoritarian decisions over alternative decision-rules, always distanced himself from the theorem as “less than trustworthy” (Estlund 1997: 189) and more recently as entirely “irrelevant” (2007: Chapter 12).

I am myself a supporter of the relevance of the CJT for democratic theory and the justification of majority rule (see Landemore ms 2007 and Landemore 2009). I have defended what I call “political cognitivism” or the position according to which it generally makes sense, among other things, to speak of the probability of the correctness of a voter’s opinion. I have also defended the empirical relevance and plausibility of the Enlightenment assumption, at least for specific uses of majority rule (such as Luppia’s

¹¹ Essentially because of the assumption that each individual has to have a greater than fifty percent chance of being right on any yes or no question (he seems unaware that the theorem in its modern form requires only that this requirement apply to the average voter, not each of them, allowing for a greater distribution of epistemic competence). Christiano also strangely categorizes the epistemic case for democracy as a subset of the intrinsic justification for democracy based on an appeal to the value of liberty.

“big choices”¹² (Lupia 2006), as well as the Sincere Voting assumption, at least in the context of mass elections since when strategic interactions vanish (with large numbers), it is not irrational to vote sincerely. I find more difficulties, as other commentators, with the Independence assumption, although I do not quite see why it could not be interpreted as an oversimplification of something plausible like the negative correlations that Hong and Page have in their own models (more on which shortly). Before I turn to alternative accounts of the epistemic properties of majority rule, let me point out an undeniable asymmetry in the way democratic theorists have received and used mathematical theorems in relation to normative justifications for majority rule.

There are at least three theorems directly related to majority rule: the Condorcet Jury Theorem, Arrow’s Impossibility Theorem, and May’s Theorem.¹³ The oldest one, the CJT (1785), is paradoxically the one who has had the least normative impact on democratic theory. Arrow’s impossibility theorem, by contrast, which is yet but a spin-off of another finding by Condorcet when he considered the case of majority rule applied to more than two options (the Condorcet Paradox of cycling majorities), has had devastating normative implications, allowing Riker to conclude that majoritarian democracy is impossible, meaningless, and inaccurate and leading generations of political scientists to repeat the mantra that “the only voting method that isn’t flawed is a

¹² “Big choices” have the following three properties : 1) they are generally binary (yes or no, or a choice between two candidates), 2) simple and widely available pieces of information (or cues) help voters vote as they would have if they knew more about the choices in question, and 3) their salience makes it easier for ordinary people to use cues efficiently. Given these three properties, Lupia concludes, big choices stack the deck in favor of competent voting, even by people who cannot answer common survey questions about politics. Lupia further suggests that it is probable that ordinary people are competent in a wider range of circumstances than many critics presume.

¹³ There is also the less well-known Gibbard-Satterwait theorem (1972), which shows that any voting method that is completely strategy-free must be either dictatorial or non-deterministic. One constraint that a social choice function must satisfy in the framework of the Gibbard-Satterthwaite theorem is the same “unlimited domain” of preferences that is also present in the framework of Arrow’s theorem so I refer to the critiques of that assumption for further discussion (footnote 2). I will say nothing here about that theorem, which does not have the prominence in democratic theory of the other three.

dictatorship” (Riker 1982 REF). Whereas the CJT is often dismissed as “too abstract” to be relevant to real-life democracy, Arrow’s impossibility theorem keeps serving as the cornerstone of political scientists’ defense of elitist democracy.

Yet if an assumption as abstract as “unlimited domain of preferences” (one of the conditions for Arrow’s theorem to hold¹⁴) makes sense, it is hard to understand why the assumption that people’s votes are statistically independent of each other (the Independence assumption of the CJT) couldn’t. Notice too that if majority rule is doomed by Arrow’s impossibility theorems, then one wonders how elitist democrats justify the rule of representatives themselves, since presumably elites must resort to some form of majority rule to make decisions among themselves and for more than two options the problems of cyclicity and strategic manipulation occur in smaller groups as in large ones.¹⁵ The suspicion is that there is an antidemocratic bias at play in the case of people who endorse the negative implications of Arrow’s theorem but reject the positive implications of the CJT.

The comparison between the use of May’s theorem and the CJT is also interesting. May’s theorem, like the CJT, only applies for a choice between two options, so the comparison is even more straightforward than the comparison between the CJT and Arrow’s impossibility theorem. Since both theorems have positive normative implications for majority rule, one might think that if democratic theorists embrace May’s theorem, they should embrace the CJT as well. Yet this is far from being the case. Since the reason cannot be found in a veiled elitism, one must turn to another

¹⁴ The other criteria are non-imposition, non-dictatorship, monotonicity, and independence of irrelevant alternatives.

¹⁵ And yet it seems that elitist democrats believe that Arrow’s theorem ceases to apply when majority rule is used by the few as opposed to the many.

explanation, which I found convincingly spelled out by David Estlund in his observation that democratic theory has been characterized in the last thirty years by a general “flight from substance” (Estlund 2007: REF). Let me explain. Since the deliberative turn, in the 1990s, democratic theorists have been focusing essentially on the intrinsic properties of democratic practices, whether majority rule or deliberation, and have stayed away from concerns about the “substance” of the choices reached through those practices. According to Estlund, this is because theorists fear that shifting the focus away from the fairness of the procedure to the question of the rightness of the outcome risks backfiring and potentially justifying antidemocratic conclusions. If one starts caring about the quality of collective decisions, what guarantees that the conclusions are going to be democratic? Haven’t epistemic considerations always served to justify the rule of an elite of knowers (an epistocracy, as Estlund calls it)? A similar logic might explain the fact that democratic theorists welcome May’s theorem, which reinforces their procedural case for majority rule on the basis of equality and fairness, but shun the CJT and its instrumental case for majority rule for fear of opening the Pandora box of possibly antidemocratic conclusions.

Even if the fear that an instrumental concern for outcome could fireback and produce antidemocratic conclusions was justified, however, that does not make it an argument. Furthermore, there are actually good reasons to believe that an instrumental concern for good outcomes actually supports the epistemic superiority of majority rule over rule of the few.¹⁶ The Condorcet Jury Theorem might not be the best argument to

¹⁶ I have myself argued that in politics, where answers are not obvious and no one can claim to know best with absolute certainty, the most reliable “knower” is not to be found in any individual or subset of individuals, but in the group itself (see Landemore ms 2007). In other words, it is safer to bet on the wisdom of crowds, or collective wisdom, than on the wisdom of any individual within the crowd.

support that epistemic case for majority rule, but as we will now see there are at least two other options: the Miracle of Aggregation and Hong and Page model of the wisdom of crowds based on “cognitive diversity.”

2. Miracle of aggregation and the Cognitive Diversity Model

The Miracle of Aggregation, which has been recently popularized by Caplan’s fierce critique of it (Caplan 2007), is a variety of statistical account of the phenomenon known as “the Wisdom of Crowds.” According to at least one version of the “Miracle of Aggregation,” collective predictive accuracy simply depends on extracting the information held by an informed elite (at least one person) from the mass of “noise” represented by other people’s opinions. As long as one person in the crowd knows the right answer and all the others make mistakes that cancel each other out, the right answer is still, so to speak, going to rise to the surface.¹⁷

Applied to the case of voting, the Miracle of Aggregation says that in the aggregate mistaken votes cancel each other out, leaving only the right answer—given by the median voter—to determine the group’s collective outcome. To go back to the example developed in the introduction, let us say that the right direction is North. Only one person in the group actually knows this. All the other choose randomly, which results in an equal split of those voters between the option North and the option North West. If the group resorts to majority rule, then the knowing voter will tip the collective choice in favor of the right answer. This is a much better option than flipping a coin, where there

¹⁷ I distinguish between an elitist, democratic, and distributed version of the Miracle in Landemore 2009.

would be only a 50% chance of getting the right answer. It is also better than letting one randomly chosen member of the group make the decision, since the probability that the knowing person would be randomly picked $1/n$, where n is the number of members in the group.

Unlike the CJT, the Miracle of Aggregation is minimally demanding in terms of the average competence of citizens, since in theory a single enlightened vote could still make the group smart, assuming that the other votes are random (or, in other words, that the mean of such aggregated votes is zero). The Miracle of Aggregation saves us, in other words, from having to assume that the average voter is more likely than not to get it right on any yes or no question (the CJT's Enlightenment Assumption).

For all its appeal, however, the Miracle of Aggregation is based on two problematic assumptions. One is the assumption that votes are randomly distributed. As Caplan (among others) rightly points out, in reality, voters are often biased in systematic ways, which would then seem to turn majority rule into a machine good at producing large-scale errors.

The other problem is that, like the CJT, the miracle of aggregation assumes that voters' predictions (votes) are independent of each other. Whatever problems plagued that assumption in the case of the CJT, they also plague the Miracle of Aggregation. We need a model of collective decision-making that is more realistic in terms of the ways people make up their minds and vote, which implies accounting for the fact that votes are often statistically correlated, not independent (i.e., the probability that voter A votes correctly depends on the probability that voter B votes correctly).

Lu Hong and Scott Page's model of the collective wisdom of group decision purports to offer such an alternative account of the wisdom of crowds that can be applied to majority rule. While Hong and Page's account is compatible with the statistical models (Condorcet Jury Theorem and Miracle of Aggregation), it translates more satisfyingly into plausible empirical conditions. Roughly, the epistemic argument for majority rule that can be extracted from Hong and Page's model is that in order to maximize our chances of picking the better of two options, we are better off taking the median answer of a sufficiently cognitively diverse group of people than letting a randomly selected individual in that group make the choice for us. This is so because, for a given group of people using different predictive models, the predictions will be negatively correlated and mistakes will cancel each other not randomly but systematically. As a result, the average mistake of the group will be less than the average mistake of a randomly selected individual, and in fact all the lesser as the difference between the predictive models used by those individuals is greater (i.e., as there is more cognitive diversity in the group).

Let me unpack that claim a little bit, and use an example. The emphasis in Hong and Page's account is not so much on the existence of a large number of votes as the cognitive diversity that those numbers can bring. Cognitive diversity is, roughly, the fact that people make predictions based on different models of the way the world works or should be interpreted. For example, in order to make a prediction about who is the best potential president out of two candidates, some of us will base our judgment on how competent on social issues he is likely to be. Others will make a prediction based on both how fiscally conservative he is and the presumed state of the economy in the coming

years. Still other people will make a prediction based on a mix of factors: the president's charisma, the current price of oil, and the prospect that Iran obtains the nuclear bomb.

Let me go over a simplified model of how this works out (I built it after an example from Hong and Page in Page 2006: 199-205). Consider 3 voters A, B and C, who are trying to make a prediction about the competence or incompetence of several presidential candidates. To capture how these three predict as a crowd making a prediction, we assume that they vote based on their predictions. As each person predicts either a competent (C) or an incompetent candidate (I), we cannot have any ties. Let us say that A is a Democrat, so he makes predictions based on how good of a job the candidate will do on social issues. Anyone at least moderately progressive is considered competent. B is a Republican so he judges the candidate based on how fiscally responsible the candidate will be. Anyone at least moderately fiscally conservative is competent. C is an Independent and he judges on a mix of both factors. C predicts that those candidates that are at least moderately fiscally conservative and either a little or moderately socially progressive will be competent.

Let us assume that Table 0 below presents the mapping from candidates' attributes to whether the candidate would be competent (C) or incompetent (I). (I offer this as an arbitrary possible standard of the "right" answer in the choice of a candidate. Any other mapping could be imaginable).

Table 0: The candidate's attribute to competence as a president mapping

		Socially Progressive		
Fiscally conservative	Highly	Moderately	A little	Not at all

High	C	C	C	I
Moderate	C	C	C	I
Low	I	I	I	I
Not at all	I	I	I	I

Let us now consider Table 1 below, which summarizes the prediction made by voter A. Note that 10 out of A's 16 answers turn out to be right (the answers in green) when compared to the "reality" defined by Table 0.

Table 1: A's Predictive Model

Fiscally conservative	Socially Progressive			
	Highly	Moderately	A little	Not at all
High	C	C	I	I
Moderate	C	C	I	I
Low	C	C	I	I
Not at all	C	C	I	I

Table 2 below summarizes the predictions of voter B, who turns out to be right 14 times out of 16 (in orange).

Table 2: B's predictive model

Fiscally conservative	Socially Progressive			
	Highly	Moderately	Low	Not at all
High	C	C	C	C
Moderate	C	C	C	C
Low	I	I	I	I
Not at all	I	I	I	I

Table 3 (following page) summarizes the prediction by voter C, who is also right 14 times out of 16.

Table 3: C's Predictive Model

Economically conservative	Socially Progressive			
	Highly	Moderately	Low	Not at all
High	I	C	C	I
Moderate	I	C	C	I
Low	I	I	I	I
Not at all	I	I	I	I

Table 4 below summarizes the “agreement set” of A and B (which is designed so that where A and B agree, they are right).

Table 4: A and B's agreement set

Economically conservative	Socially Progressive			
	Highly	Moderately	Low	Not at all
High	C	C		
Moderate	C	C		
Low			I	I
Not at all			I	I

What happens when A and C disagree? There, C becomes the pivotal voter who determines the group's prediction and what happens is quite interesting. Looking at table 4, we see that A and B make different predictions for 8 of the boxes, the boxes not in

their agreement set. Filling in Deborah’s predictions in those 8 boxes gives the predictions from the crowd shown in table 5 below.

Table 5: The group’s prediction, using majority rule

Economically conservative	Socially Progressive			
	Highly	Moderately	Low	Not at all
High	C	C	C	I
Moderate	C	C	C	I
Low	I	I	I	I
Not at all	I	I	I	I

Table 5 summarizes the group’s prediction. Note that table 5 is exactly like Table 0.

Using majority rule, the group predicts accurately every time!

This example is of course carefully crafted to do the job, which is to demonstrate how majority rule can produce more amazing results than even the CJT or the Miracle of Aggregation would predict. As Page says, “the Law of Large Numbers [CJT] cannot get you to 100% and neither can cancelling errors” (Page 2006: 202)¹⁸.

According to Hong and Page, the reason why the aggregation of predictive models does such a great job at producing correct decisions comes from the existence of negative correlations between voters’ predictions (in the CJT or the Miracle of Aggregation, by contrast, votes are supposed to be independent). I leave it to the reader to go back to the actual mathematical demonstration of the more general theorems (Diversity Theorem and Crowd Beats Average theorem) in Page 2006 (Chapter 8). Let me however illustrate how negative correlations work in the example crafted above.

¹⁸ Technically, you do get 100% accuracy with the CJT and the Miracle of Aggregation but only at the limit case involving an infinity of voters.

Take a minute to compare A and B's prediction tables (Table 1 and 2). B is right 14/16 (or 7/8) of the time. A predicts correctly in just 8 of those 14 times where B is right (i.e., 4/7 of the time). 4/7 is less than A's actual score of 10/16 (or 5/8) of predicting correctly in general. If A's probability of predicting correctly were independent of B's probability of being correct, then A should predict correctly 5/8 of the time (which is 8.735 out of 14 cases). Since 8/14 is less than 8.735/14, this goes to show that A predicts correctly less often when B is right than would be the case if her predictions were independent of B's. In other words, A and B's predictions are negatively correlated.

The good thing about negative correlation of this type is that it guarantees that where one voter makes a mistake, another is more likely to get it right and vice-versa. In the aggregate therefore, mistakes cancel each other not randomly but systematically.

Now, one may ask, where do these negative correlations come from? They come from the fact that when looking at different candidates, voters A and B do not look at the different dimensions of a same quality (or in Page's vocabulary "perspective"), here competence for office. A focuses on competence on social issues, B focuses on competence on economic issues. This produces what Page calls "non overlapping projection interpretations," that is interpretations of the candidate's competence that do not contain any of the same variables or dimensions (in that case competence on social issue or on economic issue).¹⁹ The beauty of having such different predictive models in a group is that because of the negative correlations between predictions that they entail, the

¹⁹ Page formalizes "the Projection Property" as follows: "If two people base their predictive models on different variables from the same perspective (formally, if they use nonoverlapping projection interpretations), then the correctness of their predictions is negatively correlated for binary projections" (Page 2006: 203).

group makes even better predictions than the CJT or the Miracle of Aggregation would predict (for binary predictions).

Hong and Page's account of the logic of group intelligence seems extremely promising for an epistemic justification of majority rule, at least when majority rule is used in a group of people who make predictions based on different variables.²⁰ The superiority of Page's account over the CJT or the Miracle of Aggregation is at least two-fold. First, their account gets rid of the awkward "Independence Assumption," which rendered both the CJT and the Miracle of Aggregation somewhat unrealistic in their description of what is going on when people vote. The second advantage is that Hong and Page's model supports the epistemic reliability of majority rule used among small groups (elitist democrats should be relieved). On Hong and Page's account, you do not need to have an infinity of voters for majority rule to guarantee 100% predictive accuracy (as in the CJT). Because cognitive diversity can exist as soon as there are more than one person making the prediction, the magic can work for as small a group as three people (as in the admittedly contrived example above) and is substantially increased for any addition of a person with a sufficiently diverse predictive model to the group (whereas in the CJT model, the major payoff of majority rule is at the limit, for an infinity of people, and adding one person to the group does not make much of a difference).

²⁰ A caveat needs to be added, lest the result seem too optimistic. You cannot have an infinity of variables or dimensions associated with a given perspective (say, competence for office). As the number of voters grow very large, the number of variables that people use to make a prediction may remain proportionally quite small (on top of social and economic issues, voters may look at personal charisma and foreign policy variables but they might disregard variable such as dog type or sense of humor). To avoid positive correlations as the number of people in the crowd becomes larger, people must either use cluster interpretations or they must base their interpretations on different perspectives. The interpretation used by voter C is an example of cluster interpretation.

In my view, this account of the wisdom of crowd provides a fairly compelling epistemic argument for majority rule.²¹ Of course, this account is no more immune to the problem of systematic biases than the CJT or the Miracle of Aggregation are. If citizens share a number of wrong views—racist prejudices or the systematic biases diagnosed by Bryan Caplan in economic matters (2007)—majority rule is simply going to amplify these mistakes and make democratic decisions dumber, if anything, than the decisions that could have been reached by a randomly chosen citizen. On Hong and Page’s account, however, the risk of systematic mistakes by the crowd can only happen if the group lacks both individual predictive accuracy (people are really too dumb) and diversity in the way they make predictions. Assuming minimally sophisticated voters relative to the questions at hand (and here again, Luppia’s “big choices” strike me as the right illustration of political questions where people are sufficiently sophisticated) and a liberal society encouraging dissent and diverse thinking, however, one might argue that Caplan’s worst case scenario of a situation in which the average error is high and diversity low—the condition for the worst case scenario of an abysmally stupid majority decision—is not very plausible.²²

3. The task-specificity of majority rule as a predictive tool

²¹ For reasons of space, I do not address another fascinating result by Hong and Page, The Crowd’s Possibly Free Lunch Theorem. (Page 2006: 221). This theorem demonstrates that groups where individuals make predictions based not just on different variables of a same perspective, but multiple perspectives can on occasion be as accurate in their prediction as complex regressions by experts.

²² For a more thorough answer to that objection, see Landemore 2009.

I have just given an account of the epistemic properties of majority rule, in which majority rule is presented as a reliable way to make predictions because it beats the rule of the randomly chosen member of the group. In this section, I emphasize the specificity of majority rule as a predictive tool, by contrast with deliberation, which I argue is better suited to problem-solving.

Deliberation and majority rule are often presented as two rival institutions of democracy, with deliberative democrats seeing deliberation as more central to the concept of democracy and aggregative democrats giving that status to majority rule. Those two mechanisms are generally opposed for non epistemic reasons—in fact epistemic reasons have never been major arguments for either mechanism (for an exception regarding deliberation, see Marti 2006)—but one could imagine pitting them against each other on epistemic grounds. One may ask: which of the two decision-procedures is most epistemically reliable? Let us imagine again that the question at hand is to predict whether presidential candidates are going to be competent presidents. Let us say that one group deliberates about each of the candidates and another group merely votes (along the lines of the example presented in the previous section).

I do not have the space here to reproduce the argument in favor of the epistemic properties of deliberation here, but the key idea is that in a sufficiently diversely thinking group (people with different local optima and a common global optimum), deliberators will guide each other, given enough time, to the right answer (the global optimum) (see Landemore 2007 for an illustration). People do not even need to be very smart if cognitive diversity more than compensate for their lack of accuracy (Page's Diversity Theorem). In less formal terms, the logic of deliberation is that it allows the deployment

of “the unforced force of the better argument,” which chases away bad information and flawed arguments, to only keep the good one. In the ideal case, truth rises from the debate and puts a natural end to it: unanimity.²³

Deliberation seems to have a structural epistemic advantage over majority rule in that its dialogic and diachronic dimension allow deliberators to eliminate the bad input from consideration, whereas majority rule instantaneously averages out all types of input, the good and the bad. This would seem to mean that for a given group of people, when both procedures are at their best, one should get more out of their deliberation than out of their vote.

I would argue though that pitting deliberation and majority rule against each other along epistemic lines is not a fruitful way to look at either of those mechanisms. It is much more interesting to see them as the complementary mechanisms of a given cognitive system, in which each fulfills a distinct function.

Indeed, even under the ideal or best case scenario, deliberation remains a time-consuming decision-making procedure. The “burdens of judgment” render it probable that unanimity is hardly ever reached. If unanimity ever occurs, it is generally at the end of lengthy debates, which are costly for both the participants and the polity in general. As a result, it seems that deliberation needs to be supplemented by a more efficient decision-making procedure, such as majority rule.

²³ This can hold true, of course, only under some assumptions, such as the assumption that people exchange honest arguments and disclose all of their privately held information. The empirical failures of actual deliberation, however, are well documented. Sunstein for example emphasizes two sources of distortion: informational and social pressures, showing that if deliberative settings are not carefully constructed to avoid group think, information cascades, and other polarization phenomena systematically occur (Sunstein 2006).

So, instead of pitting deliberation and majority rule against each other as rival epistemic tools, I propose to see them as complementary aspects of a larger collective decision-mechanism. This should strike one as a relatively intuitive and perhaps even simple idea. Yet it is a point that is hardly ever made, perhaps precisely because of its utter simplicity. Let me emphasize it. Deliberation and majority rule are both decision-procedures with epistemic properties but they are not meant to perform the same task. In my view deliberation is designed to deal with new problems, where practically everything is up for grabs, from the identification and formulation of the problem to the gathering of the relevant information to the formulation of possible solutions and, in fine, the prediction of which of the identified options is best. This latter predictive function of deliberation, however, is just the tip of the deliberative iceberg. To the extent that deliberation proves clumsy at generating predictions (because even disregarding the cognitive constraints that plague actual deliberation, it takes too much time to convince everyone and reach unanimity), one is better off turning to majority rule for that specific function. In other words, once the creative and brainstorming process of generating new ideas has reached diminishing returns, there is a role for majority rule, which is not simply of being the second best of deliberation, but its necessary complement.

Whereas deliberation is thus a multi-task tool, majority rule is very specifically designed to generate collective predictions among the options pre-defined by deliberation or some other mechanism. Majority rule is, if you will, the scalpel to deliberation's Swiss knife. It does just one thing but it does it fast and with a remarkable amount of precision. At least, if the epistemic argument presented in section 2 above is right, majority rule has a better chance of identifying the right answer out of two options than a coin flip or a

randomly selected person in the group. Used sequentially and repeatedly deliberation and majority rule can thus combine into a reliable cognitive system through which a society can produce good collective decisions.

4. Two Objections

Now that I have argued for both the epistemic properties of majority rule and its task-specificity as a predictive tool, let me turn to two objections that are likely to be on the reader's mind. The first objection is to the very idea that majority rule can be a predictive tool when, arguably, there is no verification mechanism to speak of in political matters. Who is to say whether Obama was, indeed, the "right" choice? How would we know? Short of a rerun of history that would produce the counterfactual world where McCain would have been elected, it is not possible to compare the outcomes and decide that, indeed, Obama was the right choice. What is the point, then, of defending majority rule as a predictive tool when nothing can falsify the argument?

The second objection consists in saying that a more plausible way to look at majority rule is to see it as a procedure that adjudicates fairly between competing interests and/or values. When we vote for a Democrat candidate rather than a Republican candidate, it is not because we think that this candidate is "right" in any meaningful sense but simply because our interests are aligned with the policies he defends or because our values are leftist and he represents best our ideological preferences. Our Republican friend who chooses the Republican candidate does so for the same symmetric reasons, as an expression of his self-interest and/or right-wing ideological values. None of us pretend to be "right" or "wrong" in any objective sense, we simply have fundamentally different preferences, whether those are based on our interest or our ideological commitments.

Let me address the first objection first. For this, I need to go back to the example of the group of friends lost in the forest, who decide to use majority rule to reach a collective decision. It is possible to view majority rule as a predictive tool in that scenario because in such a scenario, there is an objective answer. The right direction is North, North-West or none of those two options. Furthermore, even if the group never gets to figure out what the right answer is (they starve to death before they manage to get out of the forest), independent observers can tell the story as one in which the group made a mistake. There is, in other words, a procedure-independent standard of truth. One could even imagine a scientific test of the hypothesis of majority rule as a reliable predictive tool where different groups of people using either majority rule or minority rule or some other decision rule would have to repeatedly figure out the exit of the forest (I would personally bet that on average the group that follows majority rule does best).

In politics, however, there is no such procedure-independent answer and there is no impartial observer that can tell whether the choice of such and such policies or candidate was a mistake. In politics, in other words, there is no exit.

Or is there? It is true that, in political questions, we do not have access to a procedure independent standard of correctness such as “the right direction” or “how many beans are actually in that jar.” However, given enough time, it is arguable that we can to a degree verify whether an election or a policy decision resulted in a mistake or not, even if, indeed, we do not have and never will have access to the historical counterfactual in which the rival candidate would have been elected or the rival policy followed. Politics, in other words, is not so much different from trying to guess how many beans are in the jar, except that we never get to empty the jar and count the beans

afterwards (although for an attempt at counting the political beans, see Tetlock 2005, who quite successfully comes up with ways to assess the competence of expert judgment in foreign policy).

The second objection is a variation on the previous objection. It contests that the right way to look at politics is epistemic and emphasizes instead the conflictual nature of political choices and decisions. On that view, the epistemic approach naively underestimates the role played by interests and ideological differences. Sure, the objector might say, when people are lost in a forest, they have a common goal, which is to get out of the forest. But that is a silly example. In politics, people have fundamentally different goals. When they vote, they don't express a belief about the best means to an agreed upon end (the easily derided "common good") but they express different fundamental preferences for different ends. The disagreement about abortion or gay marriage illustrates this type of fundamental rift between people that no talk of "common good" can bridge. These fundamental differences between people explains that deliberation rarely produces unanimity and it also renders the epistemic view of majority rule hopeless.

While I do not entirely discard the view of majority rule as also a way to adjudicate between conflicting interests and preferences, I disagree with the exclusive interpretation of people's votes as an expression of a "preference," whether self-interested or ideological. Such an interpretation ignores the possibility that on many questions voters choose pragmatically: they see a problem (say, an economic crisis), they compare proposed solutions, and then they vote. While considerations of self-interest and ideological commitments may play a role, I tend to think that what people want most of

all is a solution to the problem. If that means going beyond their self-interest (paying more taxes) and giving up partisan affiliations, ideological bias, even prejudice (voting Democrat for a Republican, voting for a Black candidate), then they will do it. In order to avoid the controversial nature of contemporary examples, let me turn back to the American election of 1800, which illustrates what I mean by the difference between beliefs and preferences.

According to Schoffield, in 1800, the US was facing a choice between two developments paths, one commercial and one agrarian. Those development paths were exclusive of each other and, as Schoffield puts it:

[...] in 1800, which one of the choices was “true” was hidden behind the veil of the future. Madison and Jefferson clearly believed that their agrarian vision was superior. The more information available about the consequences of the two choices, the better would be the decision of the society. As a result, the 1790s saw vigorous and intense arguments about the policy choices available: about alliance with France or Britain, about the probable growth and structure of the U.S. economy, about government debt, trade protection, and so forth. These debates could not simply be reduced to interests, but were based on the *beliefs* of the protagonists (Schoffield REF).

What Schoffield describes is a collective dilemma about which direction the country should take (Agrarian or Commercial). Naturally a lot of interests were at stake. The farmers presumably preferred the agrarian path, while the merchant may have preferred an industrial path. Similarly, a lot of ideology may have come into play. Jefferson may have preferred the agrarian future because of the civic virtues that he imagined were associated with it. Other thinkers may have preferred a commercial path that presented the virtue of opening the country to the rest of the world. All those reasons to vote seem to reflect fundamental differences between people and do not leave much room for the epistemic case for voting to be made. Schoffield’s answer, however, and mine would be that while

interest and ideological preferences may affect people's choice, it does not determine it, and in that particular case, it probably did not. Yes, the farmer would have rather gone with the candidate who represented agrarian interests. But even for that farmer, there must have been other factors to consider. Further, the causal chains connecting the choice of a given political path and one's own interest are so uncertain that self-interest becomes a weak voting guide. The interest of the farmer might actually turn out to be best served by the commercial path. Even if the farmer knew for sure that the policies of the Democratic party were going to harm him, he might reason that the commercial path would serve his children better.

Thus, against aggregative democrats for whom majority rule is merely a way to aggregate preferences and adjudicate fairly between them, I maintain that majority rule is also a way of aggregating judgments and beliefs.²⁴ In fact, I agree that as we currently use it, majority rule performs both tasks, which might explain why the actual outcomes of majority rule may do not always measure up to what the theory predicts. Aware of that problem, some authors have recently suggested letting information-markets take care of the aggregation of beliefs and keep the voting to arbitrate between fundamental values (e.g., Hanson REF and Abramowicz 2009). Delegating predictive tasks to information-markets might be a good idea, but all I am saying in that paper is that, so far, it is the task performed by majority rule in our collective decision processes.

Conclusion

²⁴ I will thus for the most part ignore in this paper the obvious objections to the previous thesis that majority rule is an epistemically reliable device. Such a claim is naturally subject to a lot of qualifiers, in part because of the on-going debate about the applicability of the CJT to real world politics and the plausibility of its assumptions (particularly the Enlightenment assumption and the Independence Assumption).

I have argued that the epistemic case for majority rule is currently underrated and underused, in spite of its actual relevance for democratic theory. In conclusion I would just like to suggest that even if non epistemic reasons such as appeals to equality, freedom, or justice are what initially triggered the use of majority rule in human societies, the reason why, from a functionalist point of view, it was kept so widely in use is probably epistemic, i.e., due to its ability to produce the right kind of outcomes. It is interesting in that respect to note that a number of social animals also resort to majority rule (or something close called “quorum sensing”), obviously not because they care about equality or freedom or justice, but because from an evolutionary point of view it works for them. My guess is thus that what really keeps the practice of majority rule flourishing, in both human and animal societies, is first and foremost the fact that it is epistemically reliable. That does not necessarily diminish the intrinsic arguments for majority rule based on the values of liberty or equality and in fact might actually give them further support. Majority rule might be valuable both for the values it embodies and the results it tends to produce.

Bibliography

- Abramowicz, Michael. 2008. *Predictocracy. Market Mechanisms for Public and Private Decision Making*. Yale University Press: New Haven and London.
- Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago: University of Chicago Press.
- Black, Duncan. 1958. *Theory of Elections and Committees*. Cambridge: Cambridge University Press.
- Caplan, Bryan. 2007. *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*. Princeton: Princeton University Press.
- Cohen, Joshua. "An Epistemic Conception of Democracy." *Ethics* 97 (1): 26-38.
- Christiano, Thomas. 1996. *The Rule of the Many. Fundamental Issues in Democratic Theory*. Boulder, Colorado: Westview Press.
- Estlund, David. 2007. *Democratic Authority. A Philosophical Framework*. Princeton: Princeton University Press.
- 1997. "[Beyond Fairness and Deliberation: The Epistemic Dimension of Democratic Authority](#)," in *Deliberative Democracy*, James Bohman and William Rehg, eds., Cambridge, MA: MIT Press.
- Dahl, Robert. 1989. *Democracy and Its Critics*. Yale: Yale University Press: 243-78.
- Goodin, Robert. 2003. *Reflective Democracy*. London: Oxford University Press.
- Grofman, Bernard and Scott L. Feld. 1988. "Rousseau's General Will: a Condorcetian Perspective." *American Political Science Review* 82: 567-76.
- Grofman, Bernard, Guillermo Owen, and Scott L. Feld. 1983. "Thirteen Theorems in search of truth," *Theory and Decision* 15, 261-78.
- Hanson, Robin. To appear. "Shall we vote on values, but bet on beliefs?." *Journal of Political Philosophy*.
- List, Christian. 2001. "Some Remarks on the Probability of Cycles." Published as "Appendix 3: An Implication of the k-option Condorcet Jury mechanism for the probability of cycles." In List and Goodin 2001 (see above).
- Lupia, Arthur. 2006. "How Elitism undermines the study of voter competence." *Critical Review* 18 (1-3): 217-232.

- Mackie, Gerry. 2003. *Democracy Defended*. Cambridge: Cambridge University Press.
- 1998. “All Men Are Liars: Is Democracy Meaningless?,” in Jon Elster (ed.) *Deliberative Democracy*. Cambridge: Cambridge University Press, 1998.
- Martí, José Luis. 2006. “The Epistemic Conception of Deliberative Democracy Defended,” in Samantha Besson and José Luis Martí, *Democracy and its Discontents. National and Post-national Challenges*, Burlington, VT, Ashgate: 27-56.
- Page, Scott. 2006. *The Difference. How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton: Princeton University Press.
- Rawls, John. 1999 [1971]. *Theory of Justice*. Cambridge, MA: Belknap Press.
- Riker, William H. 1982. *Liberalism Against Populism. A Confrontation Between the Theory of Democracy and the Theory of Social Choice*. San Francisco: W. H. Freeman and Company.
- Risse, Mathias. 2004. "Arguing for Majority Rule", *The Journal of Political Philosophy*, Vol. 12, No. 1, pp 41-64
- Schoffield, Norman (Unpublished). “Madison and the Founding of the Two-Party System.”
- Sunstein, Cass R. 2006. *Infotopia: How Many Minds Produce Knowledge*. London: Oxford University Press.
- Surowiecki, James. 2004. *The Wisdom of Crowds: Why the Many are Smarter than the Few*. New York: Doubleday.