

Democratic Reason: the Mechanisms of Collective Intelligence in
Politics

Hélène Landemore
Yale University
helene.landemore@yale.edu

Forthcoming Spring 2012 in H. Landemore and J. Elster (eds),
Collective Wisdom: Principle and Mechanisms, Cambridge,
Cambridge University Press

First presented at the International Workshop on “Collective
Wisdom: Principles and Mechanisms,” Collège de France, Paris,
May 22-23, 2008. Last revised April 2011.

Democratic Reason: the Mechanisms of Collective Intelligence in Politics¹

Hélène Landemore

Yale University

Abstract: This paper argues that democracy can be seen as a way to channel “democratic reason,” or the collective political intelligence of the many. The paper hypothesizes that two main democratic mechanisms—the practice of inclusive deliberation (in its direct and indirect versions) and the institution of majority rule with universal suffrage—combine their epistemic properties to maximize the chances that the group pick the “better” political answer within a given context and a set of values. The paper further argues that under the conditions of a liberal society, characterized among other things by sufficient cognitive diversity, these two mechanisms give democracy an epistemic edge over versions of the rule of the few.

Introduction

Traditional justifications for democracy emphasize procedural or intrinsic arguments such as those based on the ideas of freedom, equality, justice and fairness. Democracy is supposed to be a good thing because its procedures—deliberation and voting in particular—express values we care about. In this paper, I explore an alternative, instrumental argument for democracy, which defines democracy as a collective decision-method valuable in part because it channels citizens’ collective wisdom or, as I will call it, “democratic reason.” This argument takes seriously the old Aristotelian idea that “many heads are better than one” and that a party to which the many contribute is better than a party organized at the expense of one person only.² In the same vein as recent work done in philosophy (e.g., Estlund 1997 and 2008), social epistemology (e.g., Anderson 2006), political sciences (e.g., Goodin 2005 and 2008; Marti 2006, Ober 2009)

¹ This paper is based on my dissertation manuscript, see Landemore 2008.

² David Estlund is one of the first contemporary political theorists to have shifted the emphasis from the intrinsic value of the democratic decision-process to its instrumental value, arguing that the normative authority of democracy is also in part “epistemic,” that is determined by the tendency of democratic procedures to deliver “correct” or “right” results overall (Estlund 1997, 2008).

and legal philosophy (e.g., Sunstein 2006 and 2009), I thus offer and analyze theoretical reasons supporting the possibility that democracy has epistemic properties that make it an instrumentally valuable decision-mechanism. More ambitiously perhaps, I propose that, in theory at least, the epistemic properties of democracy make it superior to dictatorship and any plausible variant of the rule of the few. The boldness and originality of the hypothesis presented here comes in part from the systematic comparison between rule of one, few, and many and also from the fact that I interpret both dictatorship and oligarchy in the best possible light, granting that the idealized dictator and oligarchs pursue the good of the greatest number as opposed to their self-interest.³

While the results of a comparison between the rule of many and the rule of one should be fairly intuitive (if, indeed, many heads are better than one), the second part of the claim is more controversial. If the idea of collective intelligence or collective wisdom applies to the many, it applies to the few as well. How could a group of oligarchs not outsmart the rest of the people if the oligarchs are carefully selected? This is the very thought behind the ideal of aristocracy: the rule of the best and brightest ought to be superior to the rule of regular citizens. In this paper I surmise that this might not be the case. However counterintuitive that claim may sound, an oligarchy of even the best and brightest need not be generally smarter than the rule of the many. This is so because of the crucial role of one component of collective wisdom, namely “cognitive diversity” or the existence within the group of multiple ways to see the world and interpret it.

³ I emphasize the originality of the claim because in Estlund’s approach for example, the epistemic comparison between rule of the many and rule of the few is precluded by the requirement that any claim to authority satisfy a “general acceptability requirement.” According to Estlund, even if an oligarchy of knowers is probably more epistemically reliable than democracy, it would never pass the “general acceptability” test. The hypothesis presented here is not committed to a Rawlsian framework for legitimate political authority and pursues the comparison between all forms of rule on purely epistemic grounds.

Applying the theoretical findings of Lu Hong and Scott Page about the relative importance of cognitive diversity and individual ability for collective problem-solving and predictions (Hong and Page 2001, 2004, 2009, forthcoming 2012 and Page 2007), I argue that since an ideal oligarchy of even very smart rulers can be assumed to display less cognitive diversity, at least over the long term, than an ideal democracy (direct or indirect), the few can in general and on the long run not match the epistemic competence of many moderately smart but diversely thinking individuals. In other words, to the extent that cognitive diversity is more likely to exist in the larger than the smaller group, I argue that democracy is more conducive to collective wisdom than any version of the rule of the few.

The power of the hypothesis formulated here is that, if it is true, democracy would be preferable to oligarchy on epistemic grounds even if one could identify in advance the smartest and most virtuous individuals in a given population. On the view presented here, in other words, even assuming that we could agree on who the best and brightest are and design what would be the perfect aristocracy, we would still be better off making the decision as a group rather than delegating the group's political decisions to this subset of best and brightest. To put it yet another way, the reason to include everyone in the collective decision-process is not because we cannot know who the knowers are. The reason is that there are good theoretical reasons to believe that in politics, the best knower is the group itself.

In order to support the view that democracy is, in theory at least, epistemically superior to any version of the rule of the few, I study two of the decision-mechanisms that can be said to give democracy an epistemic edge: maximally inclusive deliberation

and majority rule under universal suffrage. I argue that, to the extent that cognitive diversity can be assumed to be correlated with the number of participants, deliberation among the many is more likely to be epistemically fruitful than deliberation among the few. I also argue that given the same assumption of cognitive diversity brought by numbers, aggregation of judgments among the many can be epistemically as good as aggregation of judgments among the few, even if one makes unrealistic assumptions about the intelligence of the few. Combined together, the epistemic properties of those democratic decision-procedures theoretically give democracy an epistemic edge over any variant of the rule of the few.

A caveat about the limitations of the argument presented here is in order. This paper deliberately operates in a simplified and idealized epistemic framework. In particular I do not dwell on the possibility of epistemic failures in deliberative settings, such as the well-documented problems of polarization, hidden profiles, informational influences, social pressures, or informational cascades.⁴ Similarly I do not address the theoretical problems raised in social theory by the case of multiple, logically connected binary decisions, to which the impossibility results on judgment aggregations apply (e.g., List and Pettit 2002, Pauly and Hees 2006). Instead, I restrict the analysis to situations where the choice is between two options only. One reason for sticking to a simplified framework is that before adding in complicating factors, I first need to be able to propose a clear and intelligible model of democratic decision-making. Second, some of the problems that arise in a more complicated framework would presumably afflict all decision-makers, whether few or many, so they are not properly discriminating in the

⁴ For an overview of these problems and references to a representative literature see Sunstein 2006: Chapter Three; for a more nuanced assessment of these problems, see Mackie 2009.

comparison attempted here. For reasons of space, I focus here on the positive side of the argument. Finally, let me emphasize that the claimed epistemic properties of democracy defended here are only probable, which is to say that the argument makes room for and recognizes the possibility of democratic mistakes. The question of the proper institutional answers to bring to democratic epistemic failures—actual and probable—are important. I will not, however, address them here.⁵

This paper has five sections. In the first section, I develop the concept of “democratic reason” as a label for the distributed collective intelligence of the people, and even, occasionally, their collective wisdom. The second section turns to deliberation as the first mechanism of democratic reason and the theoretical reasons that can be adduced for its epistemic properties. The third section briefly turns to the problem of deliberation on a large scale and the relation between representation and democratic epistemic performance. The fourth section turns to majority rule and considers successively three plausible theoretical accounts of its epistemic properties: the Condorcet Jury Theorem, the Miracle of Aggregation, and a model based on cognitive diversity (Page 2007 and this volume). The fifth section consists in a long answer to Caplan (2007)’s objection that voters are “rationally irrational,” which should make it impossible for what I call democratic reason to emerge. The conclusion recapitulates the epistemic advantages of the rule of the many over versions of the rule of few.

1. Democratic reason and epistemic competence

⁵ I do not address the problem of potentially “tyrannical majorities” either. Popular as an argument against democracy since Tocqueville first magnified it, the problem of tyrannical rulers applies to minorities as well as majorities. It raises the liberal issue, orthogonal to the epistemic question addressed in this paper, of the proper limits of any collective decision method and any form of government (see for example Rehfeld 2008: 266 for a similar and more developed point).

I define democratic reason as a certain kind of distributed collective intelligence specific to democratic politics.⁶ In what follows I will indifferently use the term “collective intelligence” and “collective wisdom,” even though the concept of wisdom is richer than the concept of intelligence, including notions of experience, time-tested knowledge, perhaps even virtue, that certainly ought to be part of democratic reason but that I do not have the space to deal with in this paper.⁷ The only diachronic aspect of intelligence that I touch upon is that introduced by the institution of representation, which creates some temporal mediation between the input of citizens and its translations into actual policies.

Democratic reason is also meant in part to contrast with the less inclusive concept of ‘public reason’ in Rawls, which works as a standard of liberal justification, and perhaps as a theory of limited government, but has little to do with the reason of the public at large, being in effect the reason of representatives, Supreme Justices and the like (Rawls 1993). The concept of democratic reason is, by comparison, meant to be maximally inclusive and not a priori limited as to its substance. Similarly, I do not place a particular emphasis on the public dimension of democratic deliberation, preferring to focus on a more open-ended definition of democratic deliberation that may or may not be public in the Rawlsian sense of filtering out certain arguments and involving only “reasonable” people.

Borrowing from the cognitive sciences and suitably modifying for my purposes the concepts of distributed intelligence and cognitive artefact,⁸ I call democratic reason

⁶ Democratic reason is in my view an instantiation of the larger phenomenon of “collective wisdom” studied in this volume.

⁷ See for a broader reflection on the notions of collective wisdom and collective intelligence, Daniel Andler, this volume.

⁸ See my introduction to this volume for a proper presentation of those concepts.

the kind of collective intelligence distributed across citizens and a certain number of institutions and practices that can be seen as specifics to democratic politics. As said earlier, in this paper, I focus on the synchronic properties of groups of one, few and many and leave aside the temporal dimension of democratic reason that makes it a variety of collective wisdom.⁹ I also propose to interpret decision-procedures such as collective deliberation and majority rule as cognitive artifacts that help individuals perform a social calculus beyond their individual abilities.

The conceptual gain of an approach to democratic institutions as part of a system for intelligent collective decision-making is that it allows us to explain how the individual citizen cognitively unburdens herself by letting others, as well as the environment, process parts of the social calculus. Against political scientists worried that individual citizens lack the capacity for self-rule, which arguably places the right of the people to self-rule on shaky foundations, the concept of democratic reason allows one to reply that what matters is not just what individuals can do on their own, but what they can do with the help of political cognitive artifacts such as inclusive deliberation and majority rule (as well as other political cognitive artifacts that I am not considering).¹⁰ Another way to say this is that democratic decision-procedures are societal ways of making collective

⁹ I do believe, though, that over time democracies are more likely to accumulate collective wisdom than oligarchies or dictatorships even though I will not defend this idea here.

¹⁰ The fact that collective intelligence can be distributed not just through space but over time as well introduces an important temporal dimension into the concept of democratic reason. Democracies should be able to learn, particularly from their own mistakes, how to immunize themselves against the worst forms of individual and collective cognitive failures. It seems to me that democracies have the capacity to learn over time, as was recently evidenced in my view with the fact that a woman could run for president in France in 2007 and a Black man in the USA in 2008.

decisions that economize on individual epistemic competence. Notice that this is a distinct argument from the traditional idea that democracies economize on virtue.¹¹

Closely related to the idea of democratic reason is the notion of epistemic competence, which is the competence one has in virtue of knowing something or having a certain kind of knowledge. I distinguish it from moral competence or virtue. Virtue ensures that a ruler wants to promote the common good rather than his or her private interest or that of a subset of the citizenry. Ideally we want rulers who are both smart and virtuous. One might be smart and morally evil—see the common good and refuse to choose it. In order to keep things simple here, I will assume that all rulers—whether one, few, or many—want the same thing, namely the common good. As already noted, by doing so I deliberately stack the deck in favor of oligarchy and dictatorship.

I also distinguish epistemic competence from the mere possession of information. Information is the raw data about political facts that citizens are supposed to need in order to be able to form enlightened judgments. The relationship between information and enlightened judgment is often reduced to a simple equation by political scientists, as if the level of information was all that mattered to establish and predict epistemic competence. It is generally assumed that to the extent that citizens do not know a set of facts—the name of their senator, the capital of Japan, which candidate supports which economic platform, the meaning of ‘welfare’ or ‘liberalism’—they are epistemically

¹¹ In the following, I will use the term “epistemic mechanism,” or “mechanism” for short, preferably to that of “cognitive artifact” but by mechanism I mean essentially the same thing as a cognitive artifact, namely a device—institution, practice, etc.—that helps us facilitate a calculus or a task. Deliberation and majority rule supplement each other in producing democratic reason. Of course to the extent that deliberation and majority rule are epistemic mechanisms available to oligarchs and even, in a degenerate form, to dictators as well, I will explain why the combination of democratic deliberation and majority rule with universal suffrage is superior to the combination of deliberation among the few and majority rule among the few.

incompetent. This assumption, however, rests on an empirical claim that is in my view not fully established.

For the sake of analytical clarity, I will therefore bracket the information question until the last section, where I deal with Caplan's critique of the Miracle of Aggregation and his more general point about the fact that democracy creates disincentives in voters to get properly informed, which is a problem if low levels of information correlate with bad choices. For now, I will simply assume that the same level of information is made available to all types of rulers (one, few, or many) be it through polls, information-markets, or boards of advisors. In my view what matters for collective epistemic competence is not so much the level of individual information as the collective level of information as well as the existence of institutions that work as mechanisms gathering and processing this information for the group. I suspect that democratic institutions do not reduce the collective level of information and I doubt that democratic citizens on average individually fall below the threshold that allows them to vote competently anyway, at least on most important issues. I should emphasize that I also draw a distinction between available information—which I assume is the same for all the regimes compared—and actually processed information and the additional, refined knowledge that this may produce—which may vary from one decision-procedure to the next.¹²

Finally, I distinguish between collective epistemic competence and individual epistemic competence. As I said earlier, collective epistemic competence might be more

¹² Hence, when I later on go on to argue that deliberation contributes to pooling and bringing out relevant information, I am not reintroducing something I bracketed out initially. I simply point out that based on an identical pool of raw data, deliberation allows for the identification of relevant knowledge and occasionally produces additional knowledge or information. I further argue that inclusive deliberation is better at doing this than less inclusive deliberation.

than just the sum of individual epistemic competences, and indeed a property emerging from the right mix of individual epistemic competence and some other factors.

Controlling for virtue and information available at the level of the group, I propose that collective epistemic competence is essentially a function of two things: individual epistemic competence and the cognitive diversity of the group.

Once we have an answer to the question of which decision-procedure has, in theory, the higher collective epistemic competence given equal amount of information and virtue in the rulers, it will be easier to reintroduce the (raw) information and virtue components and see whether doing so modifies the conclusions reached.

2. Deliberation: the Forceless Force of the Better Argument

The first mechanism that arguably makes democracy an epistemically reliable collective decision-rule is inclusive deliberation, i.e., deliberation that involves, directly or indirectly, all the members of the group.¹³ In what follows I embrace a relatively classical notion of deliberation as an exchange of arguments for or against something (Aristotle, *Rhetoric*, I, 2). I also follow many contemporary deliberative democrats in adding to that definition the goal of a rational agreement or consensus on the better answer or argument (e.g., Cohen 1989: 21 and Thompson 2008; for a dissenting view though, see Mansbridge et alii 2010). Similarly, I embrace the distinction between deliberation and bargaining or negotiating (e.g., Elster 1986). On my view, deliberation, whether it is used among democrats or oligarchs, is not supposed to involve threats, promises, sophistry, or any form of ‘strategic’ rather than ‘communicative’ action. The better argument is supposed to triumph through what Habermas famously calls its “forceless force,” that is, its

¹³ I leave outside of the discussion of this paper the boundary question (of who belongs to the group), even though it is a central question for democratic theory.

obvious epistemic superiority.¹⁴ The ideal of deliberation I have in mind is characterized by its emphasis on arguments and the search for a consensus of epistemically high quality for the sake of an impending decision.

Deliberation is generally credited for three properties arguably conducive to its epistemic properties. Deliberation is supposed to:

- 1) Enlarge the pools of ideas and information
- 2) Weed out the good arguments from the bad
- 3) Lead to a consensus on the “better” or more “reasonable” solution.

The first effect would seem to be irrelevant to the comparison attempted here between ideal competing decision-making procedures (rule of one, few and many) to the extent that I have assumed an initially identical pool of information. Notice though that I did not assume that this identical pool amounts to full information. What I will be asking in this section, therefore, is whether deliberation, used among few or many, can enlarge that initial pool of information by refining the available raw data or producing new knowledge on its basis. The other two effects, I suppose, are self-explanatory.

In order to illustrate such alleged effects of deliberation, let me consider two stylized situations of what occurs in a deliberative process. I borrow the first example from the film “Twelve Angry Men.” The beauty of that example, among other things, is that it easily lends itself to an epistemic reading, since the jury deliberation assumes a procedure-independent standard of truth to be figured out: the defendant is either guilty or not guilty.

In the film, one brave dissenting jury member—number 8, played by the actor Henry Fonda—manages to persuade the other 11 jurors to reconsider the guilty sentence

¹⁴ I summarize here a huge literature on the subject, following a recent survey by Martí 2006.

they are about to pass on a young man charged with murder. Asking the other jurors to “talk it out” before making up their mind, juror number 8 takes the group on a long deliberative journey, which ultimately ends in unanimous acquittal. “Twelve Angry Men” can be seen, in my view, as illustrating the phenomenon of collective intelligence emerging from deliberation. Juror number 8, left to his own devices, would have been unable to demonstrate that the sentence was beyond reasonable doubt. Only by harnessing the intelligence of the other members, including against their own passions and prejudice, does the group ultimately reach the truth.

The contributions vary and complement each other: juror number 5, a young man from a violent slum, notices that the suspect could not possibly have stabbed his victim with a switch-blade. The perspective of juror number 5 is not only unique (no other juror was acquainted with the proper way to use a switch-blade), it is crucial to the progress of the group’s reasoning, putting in doubt the validity of a key eye-witness report. Juror number 9, an old man, then questions the plausibility of the time it took another key witness (an invalid) to limp across his room and reach the door just in time to cross the murderer’s path as he fled the building. He too contributes to changing the collective perspective on the way the crime took place. One of the hardest jurors to convince, a stock broker left unmoved by any of the other arguments, finally has to admit that a near-sighted woman is not credible when she pretends to have seen the murderer from her apartment across the street, through the windows of a passing subway, while she was lying in bed, most likely without her glasses. The deliberation process in this scenario nicely idealizes real-life deliberative processes in which participants contribute a

perspective, an argument, an idea, or a piece of information and the group can reach a conclusion that no individual by himself could have reached.

Notice that in this scenario deliberation among several people has the three properties of good deliberation. Deliberation enlarged the pool of information and ideas for all jurors, bringing to the surface knowledge about the proper use of a switch-blade and a contradiction between this proper use and the description by the visual witness of the way the victim was supposedly stabbed. Deliberation also brought to the surface a fact that many in the group had noticed—the red marks on the sides of the nose of the woman who claims to have witnessed the murder from her room—but did not know how to interpret or use. Here the proper interpretation of the fact was that the witness wears glasses, is most likely near-sighted, and the conclusion that this fact leads to is that the testimony cannot be trusted.

Deliberation also allowed the group to weed out the good arguments from the bad. Once they reach the conclusion that the visual witness is short-sighted, knowing that she reports having witnessed the murder while lying in bed, what is most likely: that she was, or wasn't wearing her glasses? Even the most stubborn juror has to admit that the argument that she was not wearing her glasses is stronger than the argument according to which she was wearing them.

Finally, deliberation in this example leads to a unanimous consensus on the “better” answer, namely the decision to consider the young convict “not guilty” given the doubts raised by deliberation.

Now let us turn to an even more stylized situation, which should bring out the logic of collective intelligence in deliberation even more clearly.¹⁵ Imagine that the French government is choosing a city to experiment with a new program. Three députés are deliberating, one from Calvados, one from Pas de Calais, one from Corrèze. They have the following respective local optima (by which I mean their subjective rankings of the options of which they are aware) with the value between parentheses being the objective value of the city for the experiment on a scale from 0 to 10.

Calvados: (Marseille (7), Caen (10))

Corrèze: (Paris (8), Grenoble (9), Caen (10))

Pas de Calais: (Grenoble (9), Caen (10))

Let me add the further assumption that each député has a higher probability of getting stuck at his lowest optimum than at his highest one. Thus, even though Caen is the better choice, the député from Calvados is not likely to think of it first, because he thinks that only big cities like Marseille will work, or perhaps because he is subconsciously prevented from choosing the capital of his own département. Similarly, suppose that the député from Corrèze is pushing Paris, which has a value of 8, over his other two local optima Grenoble and Caen, and that the député from Pas de Calais is pushing Grenoble 9 over his other optimum Caen. For whatever reasons, none of the députés thinks of his highest optimum first. Here is where deliberation can help.

The député from Calvados might start saying: “This program should be implemented in a big city so I say “Marseille (7).” The député from Corrèze says: “Good

¹⁵ Thank you to Scott Page for helping me with this example.

idea, but then Paris (8) is better.” The député from Calvados has to agree, forceless force of the better argument obliging. Then the député from Pas de Calais interjects: “Actually Paris is really expensive for the project, we would be better off applying it in a moderate sized city, which will be just as good a testbed. How about Grenoble? (9).” The député from Corrèze agrees, but the senator from Calvados then says: “Fine, but as far as moderate sized cities go, Caen (10) is even better than Grenoble (9), and less polluted too.” In the end, they can only end up at Caen.

Here again deliberation among several people has the three properties of good deliberation. The pool of information was enlarged, as the député from Calvados, who only knew about two local peaks (Marseille and Caen), ends up knowing about the qualities of Paris and Grenoble as well. The député from Corrèze learns about one other local peak (Marseille) and the député from Pas de Calais about two others (Marseille and Paris). Notice that even if the information gained is sometimes of lesser objective quality than the one the person already held, nonetheless, only by acquiring it can the members of the group reach the highest local optimum with certainty. The député of Calvados might never have considered an option he knew about, Caen (10), if he had not been spurred away from his initial choice (of value 7) by the other two députés who offered yet still suboptimal solutions (of respective values 8 and 9).

Deliberation also allowed the group to weed out the good arguments from the bad. While it seemed at first a good argument to look for a big city (Marseille, Paris), it turns out that it was better to look into moderate sized cities (Grenoble, Caen).

Finally, deliberation did lead to a consensus on the ‘best’ solution, namely the solution that allowed the group to reach the optimum of 10, when the pre-deliberative beliefs about the best solution could have been respectively 7, 8, and 9.

According to Lu Hong and Scott Page’s results on the components of collective intelligence (Hong and Page 2001 and 2004 and Page 2007), what matters most to the quality of collective problem-solving of the type described in the previous two examples is “cognitive diversity.” Cognitive diversity is the difference in the way people will approach a problem or a question. It denotes more specifically a diversity of perspectives (the way of representing situations and problems), diversity of interpretations (the way of categorizing or partitioning perspectives), diversity of heuristics (the way of generating solutions to problems), and diversity of predictive models (the way of inferring cause and effect) (Page 2007: 7). Cognitive diversity is not diversity of values or goals, which would actually harm the collective effort to solve a problem. Because of the importance of cognitive diversity thus defined, given four specific conditions, “a randomly selected collection of problem solvers outperforms a collection of the best individual problem solvers” (Page 2007: 163).¹⁶

¹⁶ The four conditions are fairly reasonable. The first one requires that the problem be difficult enough, since we do not need a group to solve easy problems. The second condition requires that all problem solvers are relatively smart. In other words, the members of the group must have local optima that are not too low otherwise the group would get stuck far from the global optimum. The third condition simply assumes a diversity of local optima such that the intersection of the problem-solvers’ local optima contains only the global optimum. Finally, the fourth condition requires that the initial population from which the problem solvers are picked must be large and the collection of problem solvers working together must contain more than a handful of problem solvers. This assumption ensures that the randomly picked collection of problem-solvers in the larger pool is diverse and in particular more cognitively diverse than a collection of the best of the larger pool—which would not necessarily be the case for too small a pool relative to the size of the subset of randomly chosen problem-solvers or for too small a subset of problem-solvers in absolute terms. Notice that the first part of this fourth condition can be thought of as Madison’s requirement in Federalist 10 that the pool of candidates to the position of representatives be large enough. For more on this, see Page 2007: 159-162.

The general point illustrated by these two examples is that it is often better to have a group of cognitively diverse people than a group of very smart people who think alike. This is so because whereas very smart people sharing local optima will tend to get stuck quickly on their highest local common optimum, a more cognitively diverse group has the possibility of guiding each other beyond that local optimum towards the global optimum. We can imagine that, in the scenario of Twelve Angry Men, if the jury had been composed of clones of juror number 8, the smartest person in the lot, they might have been stuck with the initial suspicion but unable to turn it into the firm conviction of “not-guilty” reached by the group. Similarly if all three députés were thinking exactly alike—say like the député of Calvados who thinks of his lower local optimum first—no matter how long they deliberated, their group would stay stuck on the local optimum of Grenoble (9) and would never be able to reach the higher local optimum of Caen (10). If all thought like the député of Calvados or Corrèze, they would still have a given probability of reaching the global optimum, but not the certainty of the deliberating group described above.¹⁷

Deliberation, however, is not by itself democratic. In effect deliberation can theoretically occur within one person (degenerate case) or among a few oligarchs. The two examples I gave occur among 12 people or less. What is the gain of involving large numbers? Further, isn't there a point beyond which large numbers can worsen the quality of deliberative outcomes?

¹⁷ Each member is defined by a set of local optima and a probability of getting stuck at each of his local optima. So if the deliberating group is made up of the exact same people who have a non zero probability of getting stuck at the non global optimum, the group probability of finding the global optimum might be higher than that of any individual in the group, but it won't be a 100%.

The gain of involving large numbers is that it automatically ensures greater cognitive diversity. In that sense more is smarter.¹⁸ One is thus tempted to generalize Scott Page “Diversity Trumps Ability Theorem” into a “Numbers Trumps Ability Theorem,” by which what matters most to the collective intelligence of a problem-solving group is not so much individual ability as the number of people in the group. Thus if three députés are more cognitively diverse and thus smarter than just one, then 500 should be even more cognitively diverse, and thus smarter, than three. Similarly, if 12 jurors are smarter than one, then so would 41 or 123 jurors. Of course, this assumption that cognitive diversity correlates with numbers will not always be verified but it is more plausible than the reverse assumption.

A crucial problem, however, which might dampen our enthusiasm for numbers, is a question of threshold. Deliberation involving all members of the group is not always feasible (for a defense of a national holiday that would allow the entire nation to deliberate once a year across smaller subgroups though, see James Fishkin and Bruce Ackerman’s Deliberation Day proposal (Fishkin and Ackerman 2004)). In practice, past a certain numerical threshold, deliberation turns into a chaotic mess, in which case the epistemic superiority seems to go by default to deliberation involving a smaller number of people, preferably the smarter or more educated ones.

This is where the institutional device of political representation comes into play. Representation allows the indirect or mediated involvement of the many in a decision

¹⁸ Notice that to the extent that (and if it is the case that) cognitive diversity is correlated with other forms of diversity, such as gender or ethnic diversity, the argument suggests that positive discrimination is not just a good thing on fairness grounds but also for epistemic reasons. I will not enter that complicated debate here but it is clearly one of the potential implications of an argument advocating the epistemic properties of cognitive diversity (for a defense of cognitive diversity as being in fact the “only” reason to support affirmative action, see the conclusions of the French sociologist Sabbagh 2003).

taken by the few. In other words, representation makes democratic decision-making possible when numbers are too large in a mediated way. Let me for now turn to a detailed analysis of how and under which conditions representation allows the reconciliation of the manageability of a small assembly and the cognitive diversity of a large one.

3. Representation as a projection of cognitive diversity on a small scale

I propose in what follows a non-orthodox reading of representation as an institutional device allowing for democratic deliberation on a feasible scale while preserving at least some of the cognitive diversity characteristic of the group in its entirety. I do not make the optimistic elitist hypothesis that representation ensures rule by an aristocracy of merits and talents, i.e., an oligarchy of virtuous knowers.¹⁹ In order to defend this democratic view of representatives as simply reproducing on a smaller scale the cognitive diversity of the larger group, I need to establish two things. One is that an assembly of representatives is actually distinct from a group of oligarchs. Second, I need to show why an assembly of representatives can in theory be epistemically superior to an assembly of oligarchs.

According to Bernard Manin's historical interpretation, periodic elections and accountability are the two principles that ensure the democratic nature of representatives' functions (Manin 1996). First, representatives are distinct from an oligarchy because they are elected to the position of decision-makers and legislators, as opposed to born into it, like aristocrats, or appointed by one or a few persons only, as might be the case of non-

¹⁹ It might have been, and still historically be the case, that election of representatives are seen as a way to select the "most virtuous" and "smartest" but one could also see the function that representatives are fulfilling as that of reproducing the cognitive diversity of the larger group on a smaller scale rather than elevating the average individual ability of the decision-makers.

elected magistrates or experts. Representatives also differ from a class of oligarchs, because they are held accountable to the people not only at the end of their mandate, but, one might say, throughout, in informal ways. They are under the scrutiny of their constituents who can write, call, and criticize them in other ways than voting them out.²⁰

Of course, in order for this descriptive account to be normatively compelling, one has to assume that elections are not meant, as they unfortunately tend to do in practice, to select people on criteria that only give chances to the more educated and/or the richest members of society.²¹ Otherwise the representatives would not be so different from a class of oligarchs and other selection mechanisms may become more appealing (for example lotteries²²). Similarly, one has to assume that the periodicity of the elections do in fact foster a certain turn-over of elites so that the same persons do not remain in power for too long periods of time. As to the claim that representatives are accountable, this must be combined with a properly democratic (as opposed to Burkean) understanding of their function (e.g., Urbinati 2006). Representatives are supposed to take into account their constituents' interests and judgments, not act and decide entirely on their own. Without advocating bound mandates, one can thus see the representatives' judgment as being regularly checked against the opinion of their constituency and factoring the latter

²⁰ I thus do not mean here the possibility of representatives being liable to be recalled at any time but simply the internalized pressure, in representatives' behaviors, to act as if they were constantly under the public eye and could be held accountable to the people at any time.

²¹ For a compelling critique of and solution to the problems of representative democracy in America, see O'Leary 2006.

²² See for example, Elster 1989: 78-103, Mulgan 1984: 539-560, Goodwin 1992, Duxbury 1999, Stone 2007 and Sintomer 2007. Notice that not only are lotteries arguably more just and representative than existing election mechanisms, but while they would for sure not elevate the level of individual ability--by definition the expected individual ability of those selected would be average--they would preserve the cognitive diversity of the group. Another question is, if we stick with elections, what kind of (s)election mechanism is best conducive to cognitive diversity. In selecting, say, a hundred representatives, a system of proportional representation may produce more cognitive diversity than majority voting in single-member districts. This invites an epistemic comparison between alternative democratic selection mechanisms, some of which can produce more cognitive diversity with fewer additional members. See Elster this volume.

in.²³ All in all, the idea is that, on a genuinely democratic normative ideal of representation, representatives are not meant to be an immutable elite of decision-makers, the way the best oligarchs ought to be.

Now, how does a group of representatives theoretically compare, epistemically speaking, with an equivalent group of oligarchs? Compare the democratic solution of an elected assembly of, say, 500 congressmen with the oligarchic solution of 500 individuals. Those numbers approximate the reality of the Republic of Venice in the 15th century, which was governed by a few hundreds of aristocrats. Historically it is doubtful that oligarchies were made up of the best and brightest. Imagine for the sake of the argument that the 500 oligarchs of our examples are extremely smart and knowledgeable, as well as virtuous, and, on top of this, cognitively diverse as a group. It might then seem that an oligarchy of 500 such individuals is likely to be smarter than an assembly made up of individuals chosen by regular citizens. The problem is that such ideal circumstances, regardless of how implausible they are initially, could not be maintained over time, for at least two reasons. One is that absent periodic renewal of their members, the group of oligarchs is stuck with a given level and type of cognitive diversity. Second, absent democratic accountability, the group of oligarchs have no incentive to inform themselves about the larger, changing cognitive diversity of the larger group. In effect, no matter how smart the group of oligarchs is at the beginning, it is unlikely to remain so over time. The group of oligarchs may be characterized by high political IQ, if you will, but ultimately not enough cognitive diversity.

²³ Nadia Urbinati puts it in eloquent terms, arguing that, on her normative conception, and contra elitist definitions that pit against each other representation and participation (if not representation and democracy altogether), “representative democracy... is intrinsically, and necessarily, intertwined with participation and the informal expression of “popular will”” (Urbinati 2006: 10).

In the case of representative democracy, cognitive diversity of the assembly is preserved over the long run thanks to periodic elections that renew the pool of members. Further, an elected and accountable Parliament, which is at least minimally shaped by a larger public opinion, is more likely to stay cognitively diverse than a body of oligarchs that can only count on the discipline of its members to avoid the trap of “group think,” self-serving biases, and isolation from popular input. So while there might be times when a large enough oligarchy might temporarily epistemically be equal, and perhaps even superior to, a democratic assembly of representatives, I think that over the long run this is highly implausible.²⁴

On that reading of representation, the epistemic argument for deliberation among the many presented earlier translate to representative democracy as well, provided representation reproduces effectively the cognitive diversity of the larger group in the smaller one. The claim remains the same: deliberation involving the many, in a direct form (where feasible) or an in-direct form (through representation—regardless of the selection mechanism) is superior to deliberation among the few, because to the extent that cognitive diversity is correlated with numbers, and provided that citizens are at least moderately smart on average, the more numerous the deliberating group, the smarter.

4. Majority Rule With Universal Suffrage

Deliberation is far from being a perfect or complete decision-mechanism, in part because it is time-consuming and rarely produces unanimity. In most cases, it needs to be

²⁴ A contemporary example of an oligarchy whose epistemic success arguably compares with those of democratic regimes, at least as far as economic policies and public education are concerned, would be the Communist regime in China. Because the label “communist” in fact now covers ideological positions ranging from the far right to the far left—the only common ideology being nationalism—one can argue that the policies pursued in China compare to those that would be produced by the (democratic) rule of the median voter. I owe this provocative suggestion to Pasquale Pasquino.

supplemented by another decision-procedure: majority rule. While majority rule is more efficient time-wise, it does not allow solving problems. It allows, however, choosing between pre-determined options, ideally defined in the deliberation period. I argue in this section that far from just being a fair way to settle disagreement about the choice of an option, majority rule is also a reliable way to improve the chances of the group picking the right one, where the “right” one is simply the better one compared to the other options. Majority rule aggregates individuals’ judgments about the best course of action to take or the right candidate to elect. In other words, majority rule is not only a fair way to settle on a decision when time is running out for deliberation, but a way to turn imperfect individual predictions into accurate collective ones. Again, since majority rule is available to the lone tyrant, who is the majority by himself, and a group of oligarchs, I will further need to consider whether majority rule under universal suffrage is superior to majority rule used by a minority within the larger group.

There exist at least three related but distinct theoretical arguments for the epistemic properties of majority rule: the Condorcet Jury Theorem, the “Miracle of Aggregation,” and Scott Page’s “The Crowd Beats the Average Law.”²⁵ I do not have time in this paper to cover them all, so I will provide only a brief explanation of the first two and focus on the last one.

4.1 The Condorcet Jury Theorem

²⁵ Adrian Vermeule (this volume) and others consider the CJT a mere variant of the Miracle of Aggregation. This might well be the case, in which case the CJT would probably correspond to what I call the democratic version of the miracle of aggregation, but to the extent that the relevant literature still treats them separately, I will address each account as autonomous.

The Condorcet Jury Theorem (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 ; 2) they vote independently of each other ; and 3) they vote sincerely or truthfully. To briefly illustrate, consider 10 voters, each of which has a .51 probability to be correct on any yes or no question. With so few numbers, a majority of 6 will have only a 40% chance of being right (assuming ties are counted as incorrect). However, if one expands the group to 1000 people, a majority of 501 hundred is almost 73% sure to be right.²⁶ If one expands the group to 10,000, a majority of 5001 is almost 100% to be right. As the number of people grows infinitely large, the majority is virtually certain to be right as long as people have just a slightly higher chance to be right than wrong on any binary question. The CJT, first formulated by the Marquis de Condorcet in 1785 and rediscovered by Duncan Black in the 1950s, has spawned many formal analyses in recent decades (e.g., to name a very few, Grofman, Owen and Feld 1983, Lhada 1992, List and Goodin 2001, and Bovens and Rabinowicz 2007). While the CJT has its advocates (e.g., Goodin 2005 and Sunstein 2009), some democratic theorists argue that the assumption that voters are on average better than random at making decisions on any political questions is far-fetched and render the theorem largely “irrelevant” (Estlund 2008: Chapter Thirteen). Another problem is the assumption of independence, which seems to imply that voters do not share any source of information or deliberate amongst each other (e.g., the controversy between Grofman and Feld 1988, Estlund, Waldron, Grofman and Feld 1989; Estlund 1994). This assumption is empirically highly implausible since in

²⁶ Notice that a majority of 501 means that at least 501 persons voted a certain way (i.e., probability of 73% represents the sum of probabilities that exactly 501, 502, 503, and so forth all the way to 1000 voted in the same way). See List 2004 for a discussion of the relevant statistical fact.

reality people do not pick up independent signals about the world but make up their minds based on a limited and highly dependent range of cues and sources of information. Even if the assumptions could be shown to be mathematical idealizations of a more complicated reality, it is true that the theorem has nothing to say about what is going on in actual judgment aggregation.

4.2 The “Miracle of Aggregation”

The ‘miracle of aggregation’ (e.g., Converse 1990, Page and Shapiro 1994, Wittman 1995, Caplan 2007) is another explanation for collective intelligence distinct from the Jury Theorem, although also involving the law of large numbers. The typical example illustrating the miracle is the weight guessing game observed by the 19th century statistician Francis Galton at a country fair, in which the average answer of 800 participants’ guesses regarding the weight of an ox once slaughtered and dressed turned out to fall within one pound of the right answer the right answer.²⁷ Many other anecdotes, recounted in both Surowiecki (2004) and Sunstein (2006), vividly illustrate the same “miracle” of group intelligence.

Unlike the CJT, the miracle of aggregation does not apply specifically to majority rule but explains why the average guess of large groups of people on matters with a factual answer tends to be uncannily accurate. It applies to majority rule only to the extent that majority rule is conceptualized as expressing the vote of the median voter. Furthermore, the miracle of aggregation generally applies to cases where different individuals submit their own individual estimates of some continuous quantity (e.g., the

²⁷ Some versions of the story present Galton as taking the median (which immunizes against the problem of extreme outliers). We will assume in what follows that the distinction mean/median does not matter in the cases that interest us.

weight of an ox). The options here are the different possible values of the quantity in question and there are thus more than two options. Nonetheless, to the extent that a continuum of options may be reduced to a choice between two ranges of values, say, it can be considered as an account of majority rule's epistemic properties.

The most established version of the "miracle of aggregation" explains it as the statistical phenomenon by which a few informed people in a group are enough to guide the group to the right average answer, as long as the mean of uninformed people's answers is zero.²⁸ Here collective intelligence actually depends on extracting the information held by an informed elite from the mass of noise represented by other people's opinions. As long as a sizeable minority in the crowd (the minority needs to be pivotal) knows the right answer and everyone else makes mistakes that cancel each other out, the right answer is still going to rise to the surface, so to speak. Applied to the experiment of Galton, this explanation would imply that several persons in the crowd knew the right answer and all the others made mistakes that cancelled each other out.²⁹

A more democratic version of the miracle of aggregation presents things slightly differently. This time everyone has an opinion that is roughly correct and the distribution of errors around each individual's blurry judgment is such that individual errors cancel each other out in the aggregate and the collective judgment is fairly accurate. In the example of the weight contest, this means that most people were not that far off the right

²⁸ This elitist version probably goes back to Berelson, Lazarsfeld, and McPhee (1954).

²⁹ This explanation is not very convincing for the case of the weight guessing contest, as Scott Page remarks (2006: 179). This inadequacy does not seem to have struck Surowiecki or Sunstein.

weight, although none of them knew it exactly. Page and Shapiro apply this model to account for the rationality of public opinion.³⁰

A third version of the miracle of aggregation assumes that the right answer is dispersed in bits and pieces amongst many people. As long as people express a judgment that contains one accurate piece of information and a random opinion about the piece of knowledge that they lack, the same logic of cancellation of random errors is still going to produce the right prediction in the aggregate. This explanation is unlikely to apply to the weight contest example but if it did, it would require that some people in the group knew the weight of the cow's tail, some other people the weight of the ears, etc., and that they randomized their guess about the other parts. On average all the pieces of information would aggregate to the right answer.

The miracle of aggregation, in its elitist, democratic, or distributed version, is an appealing way to account for the epistemic properties of majority rule. In effect, Galton himself, though not thinking very highly of democracy, was prompted by his own result to compare the gambling situation with democratic voting and to conclude that: "The result seems more creditable to the trustworthiness of democratic judgment than might have been expected" (Galton 1907: 246). For some, the miracle of aggregation is an even better explanation for collective intelligence and why possibly democracy works than the

³⁰ According to them, people have meaningful opinions surrounded by noise and aggregation across individuals produces an aggregation of those real opinions. For example, some citizens underestimate and others overestimate the benefits of immigration. "Even if individuals' responses to opinion surveys are partly random, full of measurement error, and unstable, when aggregated into a collective response—for example the percentage of people who say they favor a particular policy—the collective response may be quite meaningful and stable" (Page and Shapiro 1992). What Page and Shapiro imply, without saying it in so many words, is that the public is epistemically more knowledgeable as a whole than any of the individuals that make it up, which is why politicians are right to promote immigration policies based on the public's judgment (a reasoning extended by Page and Bouton to foreign policy as well, see Page and Bouton 2006).

more traditional explanation in terms of deliberation and the pursuit of rational consensus.³¹

Two main objections can be raised against the miracle of aggregation.³² First, one can deny the empirical plausibility of the hypothesis of random or symmetrical distribution of errors. Caplan points out that it is much more likely that people are cognitively biased in the same direction so that majority rule is going to amplify individual mistakes, not correct for them (Caplan 2007). I address that objection in the last section of this paper. More problematically, the miracle of aggregation relies on the same assumption of statistical independence of individual judgment as the CJT. While the assumption of independence may perhaps be interpreted as a mathematical idealization of what is really going in judgment aggregation and might ultimately be relaxed in some sophisticated versions of both the CJT and the Miracle (e.g., Boland, Proschang and Tong 1989 for the CJT), in what follows I consider a third account of collective intelligence that is not statistical but cognitive: it opens the black box of voters' decision-making process.

4.3 Cognitive Diversity

In a book (2007) and a series of articles with Lu Hong (2001, 2004, 2009), Scott Page proposes a different account of why large groups of people can make good judgments

³¹ Cass Sunstein for example sees it as a “Hayekian challenge to Habermas” (Sunstein 2006). In fact it is both unclear that the miracle of aggregation is the same thing as the invisible hand mechanism at work in the emergence of the prices of goods or information in markets, and that democratic deliberation is made superfluous by information aggregation through majority rule, polls or markets.

³² An additional, practical objection would also point out that majority rule generally involves a choice between discrete options, but rarely allows for the kind of “quantitative” and continuous voting observed in the ox-weight guessing game or information-markets. This is not a very powerful objection since the logic of the miracle of aggregation theoretically works even with the reduced choices offered in elections.

and, in particular, accurate predictions. Unlike commentators who dismiss Page's model as less compelling than a Deweyan account of democracy's epistemic virtues based on deliberation (Anderson 2006), I think his account provides a fine-grained and empirically plausible account of the epistemic properties of judgment aggregation, which can be used to account for the epistemic properties of majority rule. Although Page's model generally applies to numerical predictions that are not of a binary form (e.g., predicting sales figures), it can be applied to scenarios where judgments are binary as well (e.g., predicting whether a candidate is competent or incompetent). In my view, Page's model can be used as a nicely tailored account of majority rule's epistemic properties to be combined with, rather than pitted against, an account of deliberation's distinct epistemic properties.

Unlike the miracle of aggregation, Page's account does not rely on the assumption of an infinity of independent signals nor on the idea of random or symmetric mistakes canceling each other out.³³ What matters for collective intelligence is the existence of negative correlations between people's predictive models, which tends to lower the collective error and make the group smarter than the average individual within it.

Negative correlations are themselves the result of cognitive diversity in the group (Hong and Page 2009).³⁴ Cognitive diversity is, roughly, the fact that people make predictions

³³ Scott Page in fact dismisses both the Condorcet Jury Theorem and the Miracle of Aggregation because, in his view, they both implausibly presuppose that voters receive an infinity of independent signals that they pick up in order to make a prediction. In reality, he argues, people make up their minds based on a limited and highly dependent range of cues and sources of information. The infinity of signals assumption assumes more cognitive diversity than is empirically plausible. It is, according to Page, a "heroic assumption" (Page 2007: 192).

³⁴ Hong and Page 2009 demonstrate that using independent *interpretations* (not predictions) entails negatively correlated predictions. More specifically, the gist of the paper consists in demonstrating that "seeing the world independently, looking at different attributes, not only does not imply, it is inconsistent with, both conditional independence of signals and independently correct signals" (Hong and Page 2009: 18). In other words, Page's model does not require that people make uncorrelated predictions but that they

based on different models of the way the world works or should be interpreted. Cognitive diversity should be distinguished from both its symptoms (a different set of viewpoints or opinions) and its possible rootcauses (gender or ethnic diversity) as well as a diversity that is actually epistemically harmful, namely a diversity of goals or values (in an epistemic framework as in Page's cognitive model, all the members of the group are supposed to pursue the same goal and want the same thing, namely make an accurate prediction).

The good thing about negative correlations between individuals' predictive models is that they guarantee 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. The result is expressed by Page's "The Crowd Beats Average Law": the accuracy of the group's prediction is systematically better than the average accuracy of its members. In other words, the group necessarily predicts better than its average member. Further, the amount by which the group out predicts its average member increases as the group becomes more diverse (Page 2007: 197).

To illustrate briefly, consider the case of an election between two candidates, and let us assume that the point of voting in that election is, among other things, to identify who is the fittest candidate for office. Individually, each of us will make a prediction based on a limited number of factors: some of us will base our judgment on how competent with social issues a candidate 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

make up their mind independently. Except in the very implausible scenario where all reasonably informed individuals each ignore a different piece of information, their predictions will not be independent but negatively correlated.

candidate's charisma, the current price of oil, and the prospect that Iran obtains the nuclear bomb. When looking at the candidates, we will thus look at different dimensions of the same quality (or in Page's vocabulary "perspective"), which is in this case competence for office. 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 (for example competence on social or economic issues).³⁵

The beauty of having such different predictive models in a group is that, because of the negative correlations between predictions that they entail, the group makes even better predictions than the CJT or the Miracle of Aggregation would predict (for binary choices). In the example of voters deciding whether candidates are competent or incompetent, the idea is that Republicans will be more likely to be right when Democrats are more likely to be wrong and vice-versa. When a Republican and a Democrat disagree, the tie can be broken by the perspective of a third cognitive perspective, for example that of an Independent. Where the Republicans and the Democrats agree, however, the candidate is likely to be correctly described as either competent or incompetent (see Landemore 2009 for a more detailed analysis).

Let me now reformulate in more general terms the epistemic argument for majority rule that can be extracted from Hong and Page's model. The argument 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

³⁵ 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 2007: 203).

letting a randomly selected individual in that group make the choice for the group. 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 voter, and in fact all the lesser as the difference between the predictive models used by those voters increases (i.e., as there is more cognitive diversity in the group).

Hong and Page's account of the logic of group intelligence is in my view extremely promising for an epistemic justification of majority rule, at least when majority rule is used in a group of voters who make predictions based on different variables.³⁶ The superiority of Page's account over the CJT or the Miracle of Aggregation is at least two-fold. First, their account circumvents the problematic assumption of judgment independence, which rendered both the CJT and the Miracle of Aggregation somewhat unrealistic. As I understand it, the independence assumption is now applied, more plausibly, not to people's actual judgments (their outputs) but to the cognitive processes leading to those judgments (i.e., the predictive models people use to generate judgments and predictions about the world). In other words, by internalizing the independence constraint, Page's model makes it possible for citizens to share information, premises,

³⁶ 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 variables 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 someone who would judge a candidate both on his competence on social issues and his competence on fiscal issues is an example of cluster interpretation.

and even conclusions, while remaining independent in terms of the cognitive processes that treated the shared information and generated the shared conclusions.

The second advantage is that Hong and Page's model supports the epistemic reliability of majority rule used among small groups.³⁷ Unlike what happens with the CJT or the Miracle of Aggregation, you do not need to have an infinity of voters for majority rule to guarantee 100% predictive accuracy. Because cognitive diversity can exist as soon as there is 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. In the CJT, by contrast, the major payoff of majority rule is only with large numbers, and adding one person to the group does not make much of a difference.

The flip side of this, however, is that in Page's model there is a theoretical limitation to how much including more and more people improves collective judgment. Cognitive diversity in judgment aggregation is not a linear function of numbers and there are in fact diminishing returns to adding more people past a certain point. The reason, in brief, is that there are only so many different variables associated with a given perspective (competence for office for example) that people can use when trying to make a prediction. In trying to predict the most competent candidate, some people will use as a variable the candidate's personal charisma, others his ideological affiliation, and some other still his past record as a politician. As the number of voters grows very large however, the number of variables that people use to make a prediction may remain proportionally quite small. To avoid positive correlations as the number of people in the

³⁷ In fact their account is more optimistic for small groups than very large ones. I do not have the space to address this concern here. It seems to be the case though that majority rule used in representative assemblies is more likely to have epistemic properties than majority rule used in referendums.

crowd becomes larger, people must either use cluster interpretations, for example using both personal charisma and past political record as variables, or they must base their interpretations on different perspectives, for example by trying to predict a candidate's ability to win the elections rather than her competence for office. What the cognitive model suggests here is that it is probably better to aggregate the views of a limited number of representatives than those of millions of voters. At the scale of an assembly of representatives, aggregating more judgments can be expected to have increasing returns in terms of cognitive diversity, which may be lost when we aggregate the views of millions of citizens. Further analyses of the implications of Page's model are certainly necessary. Suffice to say here that his account of group competence provides, in my view, a fairly compelling epistemic argument for majority rule.³⁸

What are the implications of those considerations for the idea of democratic reason? I have argued that majority rule and, in fact, any democratic mechanism that aggregates individual judgments into collective judgments has epistemic properties. Since the group's prediction is epistemically superior to that of the average citizen in the group, we have an argument why the rule of the many is superior to the rule of one (when the one is randomly chosen). That does not give us a maximal argument for majority rule though, since majority rule among the many does not systematically beat majority rule used among a few smart people. It is the superiority of democratic deliberation over oligarchic deliberation that allows us to derive the more ambitious claim.

³⁸ For reasons of space, I do not address another fascinating result by Hong and Page, The Crowd's Possibly Free Lunch Theorem (Page 2007: 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.

5. The Problem of Voters' "Rational Irrationality" and Systematic Cognitive Biases

According to Bryan Caplan, the main problem with the optimistic conclusions about group intelligence that I have derived above is that in some way or another they rely on the assumption that there is a symmetrical distribution of errors around the right answer so that the mean of these errors is zero ("miracle of aggregation") or that errors are negatively correlated (Page's model), when they might in fact be positively correlated. In economic matters, at least, Caplan argues that voters are systematically biased in the same wrong direction. How does that argument affect the hypothesis proposed here that democracy is the epistemically superior collective decision-method?

First, empirical observations of the way American democracy functions or fails to function are not enough to falsify a more general claim about the epistemic properties of a model democracy. The empirical problems Caplan points out may be due to the fact that American democracy is not a real democracy in the sense used in this paper, lacking too many of the features I have insisted on (for example a representative system preserving the cognitive diversity of the larger group). Caplan, however, backs up his observations with a theory of the rationally irrational voter, which seems much more worrying. I will take both objections seriously and will address them in turn, starting with the theoretical objection.

The theoretical objection is powerful if one accepts Caplan's conceptualization of the voter as rationally irrational. There are, however, many problems with this theory (see Elster and Landemore 2008 for a critique). In particular, I would emphasize here that Caplan assumes in voters a form of self-interest, incompatible with the epistemic framework of this paper. Even when they vote ideologically (for the "common good") as

opposed to “rationally” (for their pocket book), as in for example, the case of the rich Hollywood actor voting for higher taxes, voters are only doing so, on Caplan’s account, because of the unlikely prospect that their votes would be pivotal. So, in effect, the voters are still first and foremost preoccupied with voting their self-interest (in the form of a warm-glow effect when the impact of their vote is too low to make a difference) rather than promoting something like the common good.³⁹ Arguably, on Caplan’s account, if their vote mattered at all, they would revert to voting their pocket-book. This perspective is utterly incompatible with the epistemic framework of the argument presented in this paper, which assumes that people are voting what they think is right for the common good, no matter how unpleasant it is for them, whether ideologically or economically. The theoretical divergence runs so deep that it is hard to see any point of intersection between those two models, which leads to drastically different predictions about the epistemic quality of democratic output.

Regardless of that theoretical divergence, however, what about the objection that even an epistemic framework may be challenged by the existence of systematic biases in voters, whether these biases come from ignorance, irrationality, or anything else? It is true that an account of collective intelligence based on cognitive diversity 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 systematic economic biases—majority rule is simply going to amplify these mistakes and make democratic decisions worse, if anything, than the decisions that could have been reached by a randomly chosen citizen. In the account of collective intelligence that I embrace,

³⁹ There is in fact an incoherence in the description of the “rational purchase of altruism” (see Elster and Landemore 2008).

however, which emphasize cognitive independence, the risk of systematic mistakes can only happen if the group lacks both individual predictive accuracy (that is, people are not sufficiently intelligent) *and* diversity in the way they make predictions. Assuming minimally sophisticated voters relative to the questions at hand and a liberal society encouraging dissent and diverse thinking, however, 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 unintelligent majority decision—is not very plausible. In other words, the possibility of systematic biases on a majority of political issues is not very plausible.

Furthermore, deliberation can play a role in the epistemic argument for democracy I have developed in this paper that it cannot play in Caplan's model. When it comes to majorities making mistakes, my argument at least allows for the possibility of self-correction over time and through the means of public deliberation, whereas Caplan, it seems, would bring in the experts or exit politics altogether (for the market). In actual democracies, it is interesting to see that where systematic biases scenarios have been historically observed to exist, on race issues for example, most changes had to come from evolving majorities themselves, through a democratic process of collective self-reflection and public deliberation.⁴⁰ Democratic deliberation, which includes the experts as welcome but non exclusive voices, is a central part of the argument made in this paper and offers a possible solution to the problem of the occasional systematic mistakes that the public can make, a solution never seriously entertained by Caplan.

⁴⁰ In that deliberation, some may want to see key Supreme Court decisions as a part of, and some others as an alternative to, the democratic dialogue, which somewhat complicates the equation. But one could always argue that the way Constitutions tie the hands of the people on some issues was itself an initially democratic decision that a people made at some point in order to protect themselves against their own predictable irrationality, by creating constitutional safeguards for minorities for example (Elster 1977).

Let me now turn to the empirical challenge based on the measurement of systematic biases in actual, American democratic citizens. Answering that challenge will take us on what may seem as somewhat of a detour in light of the larger argument. It will allow us, however, to address the larger question of the relationship between information and political epistemic competence.

Using empirical evidence borrowed from the literature on “enlightened preferences” (essentially Althaus 2003), the Survey of Americans and Economists on the Economy (SAEE),⁴¹ and the results of his own comparison between the public’s preferences and those an “enlightened public” virtually endowed with a PhD in economics, Caplan diagnoses four main misconceptions in the average American citizen with respect to economic questions: an anti-market bias, a protectionist bias, a pessimistic bias, and a job-oriented bias. Assuming that economists are right that all things equal otherwise market mechanism is a good thing, free-trade creates more riches than it destroys, growth is more likely than stagnation, and GDP increase matters more than job preservation, then the people are wrong to hold opposite views and ask for policies based on such beliefs. The problem is not solved, or solved only to an insufficient degree, by the fact that policies are made by a priori slightly more competent representatives. To the extent that representatives are held accountable to the citizens, they only have limited leeway to improve the course of things. Consequently, Caplan concludes that, on economic questions at least, we would be better off with less democratic input. He

⁴¹ The survey is based on interviews with 1,510 randomly selected members of the American public and 250 economic Ph.D.’s and designed to test for systematic lay-expert belief differences by asking questions such as whether various factors are a “major reason,” “minor reason,” or “no reason at all” why “the economy is not doing better than it is.”

himself seems to suggest more delegation to economists and, whenever possible, to markets themselves.

Here, I will raise three criticisms. The first criticism bears on the elitist premises of the book and the method used to measure citizens' incompetence. Second, even granting that Caplan is right about the economic incompetence of the average voter, I would argue that the implications for democracy are not nearly as bad as Caplan would like to suggest. Finally, I object to the alternatives implicitly offered by Caplan. It is indeed unclear that the oligarchy of experts that Caplan sometimes seems to advocate would necessarily do much better overall than a democracy. As to market mechanism, it is not in my view a political alternative to any form of government but a mere allocation tool in the hands of one, few, or many—thus leaving untouched the question of who should rule.

Let us first address the methodological question. The first benchmark of voters' bias is knowledge of objective facts. As Caplan observes, "the simplest way to test for voter bias is to ask questions with objective quantitative answers, like the share of the federal budget dedicated to national defense or Social Security" (25). Caplan, however, does not dwell on that first standard, acknowledging that "the main drawback of these studies [that measure the mastery of factual knowledge] is that many interesting questions are only answerable with a degree of ambiguity" (25). Indeed one could argue that *no interesting political question* can be answered without such a degree of ambiguity, which raises the general issue of the relevance of a great deal of public opinion research that measures the ability to answer textbook political questions. Since the standard of

objective facts reappears through the back door of the notion of “high political IQ individuals,” let’s say a few more words about why this standard is unsatisfying.

First, information is distinct from competence and the causal link between the holding of a certain type of information measured by surveys and the competence to make political choices is not easy to establish (however “intuitive” it is sometimes argued to be). In fact most existing studies (e.g., Luskin 1988, Delli Carpini and Keeter 1996) fail to demonstrate a causal link between the inability of people at answering certain types of political quizzes and their alleged political incompetence, namely the inability to make the right choices or holding the “right” policy preferences. This is so in part because the design of factual political questionnaires smacks of elitism, measuring a type of knowledge relevant for policy analysts and journalists, but not necessarily the only one conducive to smart political choices (Lupia 2006).

The difficulty of establishing a causal link between low information level and political competence comes also from the fact that it is hard to find a good empirical benchmark for political competence that would be distinct from a good benchmark for information level. The fact that educated people are good at answering political quizzes does not entail 1. that the policy preferences of the educated are better as a result (unless you take such policy-preferences as the standard but then you are begging the question), 2. that the policy-preferences of “know-nothings” or low political IQ people (as defined by such tests) are wrong.⁴² The kind of factual knowledge measured by public opinion surveys is as crude a measurement of political competence and there is no reason why the

⁴² After all, even the scenarists of the T.V. show “The West Wing” know that you can be a competent director of communication at the White House and be unable to say three correct things about the history of the White House (West Wing, episode 1).

burden of the proof should be on people who deny the connection between political IQ as it is measured by existing empirical surveys and actual political competence.

Let us now turn to the second standard: the “enlightened preferences” of a hypothetical educated public, i.e., a group of people that is demographically representative except that they are as politically knowledgeable as possible. The method used by Althaus (2003) consists in administering a survey of policy preferences combined with a test of objective political knowledge to a group, estimating individuals’ policy preferences as a function of their objective political knowledge on factual matters (e.g. how many senators each state has) and their demographics (income, race, gender), and, finally, simulating what policy preferences would look like if all members of all demographic groups had the maximum level of objective political knowledge. In other words, the goal is to compare the policy preferences of regular citizens with those of their “more educated” selves, controlling for race, gender, income and the like.

The enlightened preference approach permits testing the plausibility of the theory of the “reasoning voter,” according to which people vote roughly with little knowledge what they would vote if they had maximal information thanks to cognitive shortcuts, heuristics, and on-line processing. The major result of this approach is to show that, no, people would probably not vote the same way since they at least do not have the same preferences when they are little informed and very informed. They tend to be more socially liberal and economically conservative in the second case (Althaus 2003: 143). Scott Althaus uses the discrepancy between the public’s preferences and those of its more “enlightened” self to criticize the representativity of opinion surveys and their usefulness in assessing the public’s voice. Caplan goes one step further, using those results to

suggest that democracy itself, which follows more or less the unenlightened policy preferences of the many, is flawed.

Consider however that the definition of “enlightened preferences” hinges on a concept of education that is correlated with the ability to score well on political IQ tests (“a test of objective political knowledge”). The standard of “enlightened preferences” is thus not much different from the knowledge of objective facts (since it is highly correlated with it). But we just saw that knowledge of objective facts might well be both an elitist measure of political knowledge and potentially irrelevant to the ability to pass a political competent vote. So what this approach does is take as the standard of “enlightened” judgments preferences correlated with an elitist and possibly irrelevant form of knowledge and then argue that the discrepancy between the actual public’s preferences and those “enlightened preferences” is meaningful, and in fact an embarrassment for democracy. Such conclusions, however, merely reflect a belief present in the premise, namely that regular people are wrong and the elites right. How is that not begging the question of who has epistemic authority in the first place?

The third standard consists of the economic preferences of a simulated public that is both demographically representative and endowed with the knowledge of the holder of a PhD in economics. The key difference between Caplan’s approach and the previous approach is that “political scientists usually measure knowledge directly, while my approach proxies it using educational credentials” (55). So, in effect, whereas the second type of approach boils down more or less to using the standard of objective facts (through the notion of political IQ) to assess the public’s preferences, Caplan’s approach takes as the ultimate standard the knowledge of experts. Another difference is that the

competence that Caplan is trying to assess is slightly narrower than that measured by political scientists, since Caplan is interested only in political questions with an explicitly economic dimension, for which economic knowledge such as that measured by a PhD diploma might seem directly relevant (more so at least than “objective political knowledge” with respect to political competence).

So let us consider why the fourth standard—experts’ knowledge—is problematic. First, Caplan constantly writes as if there was no difference between questions of economics (the science) and economic questions, which are political questions with an economic dimension. Just because PhD holders in economics are the best at answering questions in the science of economics does not make them the most competent in answering political questions with an economic component (although their input is most likely of value). In fact, if you deny that economists’ political beliefs are absolute truths, the discrepancy between these beliefs and those of the public does not necessarily say much.

Despite initially acknowledging that political questions cannot be answered without a degree of ambiguity, Caplan does write as if the beliefs of economists were on a par with mathematical truths. Here is a typical example. Caplan argues that “elitist though it sounds, [inferring the existence of systematic biases in the public from the existence of systematic differences between economists and non economists] is the standard practice in the broader literature on biases” (52). Caplan goes on to appeal to the authority of no less than Kahneman and Tversky who describe their own method this way: “The presence of an error of judgment is demonstrated by comparing people’s responses either with an *established fact*... or with an *accepted rule of arithmetic, logic,*

or statistics” (52).⁴³ Caplan thus draws a clear parallel between the consensual beliefs of economists on the one hand and objective facts or the rules of arithmetic, logic or statistics on the other hand. This parallel, however, is highly misleading. To the extent that economic beliefs are about facts (the share of foreign aid in the federal budget) or about mathematical theorems, they are not necessarily relevant, or not directly so, for political decisions. To the extent that these beliefs are more “political”—even the least controversial ones, like “free trade is good” or “people are not saving enough”—they are much more contingent on a shifting cultural and possibly ideological consensus among experts than Caplan allows for. By playing on this ambiguity between pure questions of textbook economics and political questions with an economic dimension, and by misleadingly identifying the beliefs of economists at a given time with factual truth or mathematical principles, Caplan in fact begs the question of who has authority in the first place. In his view, on anything remotely economical, economists know better. If you deny that premise however, none of Caplan’s conclusions follow.

Both the “enlightened preference” approach and Caplan’s “enlightened public” approach beg the question of who is politically competent in the first place, whether it is people with a high political IQ or economists. Caplan supports that way of proceeding by arguing that “the burden of the proof should be on those who doubt the common sense assumption that we should trust the experts” (2007: 82). One might reply, however, that democracy is premised on the very rejection of that “commonsense” assumption. For

⁴³ Caplan further comments ““established” or “accepted” by whom? By experts of course” (52). Notice however that unlike mathematical truths, which can be accepted by everyone too, not just experts, economic truths are never as universally endorsed.

democrats since at least the Sophist Protagoras,⁴⁴ politics is the realm where no one is to be trusted more than others to begin with. This is why, as Socrates could observe, the Athenian Assembly behaves very differently when the problem is to build an edifice or a ship, than when the question is to figure out the good of the city. In the first scenario, the Assembly calls in architects and shipbuilders and if someone who is not considered a competent technician in the relevant field speaks up to give his opinion, the crowd boos and shames him into silence. By contrast, Socrates goes on:

When the question is an affair of state, then everybody is free to have a say—carpenter, tinker, cobbler, sailor, passenger; rich and poor, high and low—any one who likes gets up, and no one reproaches him, as in the former case, with not having learned, and having no teacher, and yet giving advice (*Protagoras* 319d).

For Athenian democrats, the real test of competence and expertise in politics is the ability to convince others in the Assembly. This is why, even if ultimately only the better arguments and information are supposed to triumph, everyone has the right to speak up.

I just criticized Caplan for begging the question of who is right in politics when defining the benchmark of competent answers as those of people who think like economists. Aren't democrats begging the question the other way around by denying that there are experts in the first place?

The positions are not exactly symmetrical. In Caplan's case, the question of who knows best and what the right answers are is a priori locked and determined. The economists know better, their answers are the right one, and thus any deviation from their position must be measured as a bias. On the democratic view, there is genuine agnosticism as to who knows best and what is the right answer, at least at the outset. Who knows best and what the right answers are can only be determined, in the short term, on

⁴⁴ According to Cynthia Farr, "Protagoras was, so far as we know, the first democratic political theorist in the history of the world" (1988: 77).

the merits of different claims competing in the public space and later, on the much longer term, by a retrospective look on how well the country did overall given that such and such policies were implemented or even by comparing expected to actual results for every chosen policy. At the moment of decision-making, when such hindsight is not available, the benchmark of right political answers is, for Caplan, whatever economists say, whereas for democrats the benchmark is only the “forceless force of the better argument” (which does not mean that the best argument will always triumph) and/or majority outcomes (which does not mean that the majority is always right).

Of course, when looking at certain resilient discrepancies between what the public thinks and what economists think, there are cases where experts are probably right and the public probably wrong. For example, people tend to think that “taxes are too high” or that “foreign aid spending is too high” (Caplan 2007: 57), whereas economists and the “enlightened public” sensibly differ. A lot hinges here on what is meant by “too high” and the point of comparison implied in these judgments so I am not entirely conceding that these are clear-cut examples of biases on the part of regular citizens. Still, one might plausibly argue that, on certain questions, an oligarchy of knowers would make more enlightened decisions than a democracy. There are, however, a few reasons why granting topical incompetence here and there does not affect the general argument for democracy developed in this paper.

First, topical incompetence does not establish global incompetence and in particular the meta-incompetence to recognize one’s topical incompetence. Even if we accept that citizens are bad at answering political questions of an economic nature, that does not mean that they are not reasonable enough, that is minimally competent, to

acknowledge that fact and accept institutional arrangements that compensate for it, such as delegation of some decisions to acknowledged experts.

Second, delegation of some choices to experts does not imply the failure of democracy. Democracies that delegate some decisions to a few unelected individuals do not ipso facto turn into oligarchies. The fact that the consent of the people was initially obtained for this delegation to take place (directly or through their representatives) still makes the decisions of those experts “democratic” in a larger sense. To the extent that the independence of central banks itself was a democratic choice, it should testify to democratic intelligence on Caplan’s view, since his story is supposedly voter-driven. Conversely, the decision-power of democratically authorized experts on some economic questions does not prove the superiority of oligarchy over democracy but simply establishes the necessity of having some efficient technocratic cogs in a larger and complex democratic structure of governance.⁴⁵ The relevant comparison for my purpose in this paper is not between democracy and that technocratic branch of the government but democracy and oligarchy when both are equipped with a competent technocracy of that kind. John Stuart Mill thought that the only virtue of a monarchy was its bureaucracy, whereas the virtue of a democracy was its bureaucracy plus the intelligence that goes into overseeing it (Considerations on representative Government, Chapter 5 and 6). Similarly, my argument leads me to conclude that when both democracy and oligarchy are equipped with a competent army of experts, democracy would still, on average and on the long term, outperform oligarchy.

⁴⁵ In fact, this voluntary delegation of technical economic questions to experts might be all that Caplan ultimately advocates.

Third, even if Caplan is right about voters' topical incompetence, particularly in economic matters, why does he not consider the possibility that such topical incompetence might be solved over time through education and public debates? I have already mentioned that deliberation might be a solution to systematic biases. But Caplan seems to equate observed ideological preferences with deeply entrenched (bad) cognitive biases and heuristics. In the same way that people are known to suffer from base-rate neglect⁴⁶ or to be subject to framing effects,⁴⁷ Caplan suggests that they are systematically anti-free-trade and pro-job security. But an anti-market or a pro-job bias is of a different nature than an inability to calculate probabilities correctly or see a glass as equally half-full and half-empty. Such economic biases are less due to the limits of human cognitive abilities and more to cultural factors. After all, while all human beings may suffer from some form of base-rate neglect, Americans are actually much less obsessed with job security than Europeans. President Clinton during his presidential campaign could thus warn the American public that "they would have to change jobs 7 to 8 times in a life-time"—a discourse utterly unthinkable in a French context. Racial and sexist prejudices have considerably diminished in most Western democracies over just a few generations.⁴⁸ These facts suggest that some biases can be corrected, at least partially. Maybe economic biases are of a more enduring nature but Caplan does not demonstrate this for a fact. Education and a more deliberative democracy, however trite that may sound, may well be the answers to (at least some) of the flaws of our existing democracy

⁴⁶ The base-rate neglect or fallacy consists in neglecting the prior probability of some hypothesis H when trying to assess the conditional probability of this hypothesis given some evidence E.

⁴⁷ They give different answers to a same question that frames things differently.

⁴⁸ The US has now a Black president.

The final objection I will raise is against an apparent implication of Caplan's indictment of democracy—that we would be better off with an oligarchy of experts—⁴⁹ is that groups of experts are not fool-proof either. Philip Tetlock showed in his study of “political judgment” that when it comes to assessing a problem and making political predictions, political “experts” do hardly better than lay people and, on the purely predictive side, are in general outperformed by simple statistical regressions. Striking what should seem like a deadly blow against the idea that politics is a matter of expertise, Tetlock concludes that it does not really matter *who* the experts are (economists or political scientists or philosophers...), nor *what* they think (ideologically, i.e., whether they tend to be pro-market or socialist). What matters is the *way* political experts think, namely whether they think as “foxes” or as “hedgehogs.”

Borrowing Isaiah Berlin's ideal-types, Tetlock characterizes “foxes” as eclectic thinkers with an ability to use different frameworks and theories. By contrast hedgehogs are dogmatic thinkers with a one-size fits all theory of the world. From what Tetlock could empirically observe, foxes are almost always better forecasters than hedgehogs. Tetlock also shows that both foxes and hedgehogs are generally outperformed by statistical regressions. If political experts—pundits, political campaign leaders, diplomats etc.—tend to overestimate their knowledge, analytical skills, and ability to predict what will happen, it is probable that economists—who tend to fit the model of the “hedgehog” or dogmatic thinker described by Tetlock—suffer from the same cognitive failures.

⁴⁹ Caplan would deny that this is the solution he advocates, yet everything, from the cover of the book to many assertions in it, invites an anti-democratic reading. Caplan could have tried harder to dissuade the reader from thinking that what he ultimately advocates is rule of the experts, in the same way as he would have liked to see Tetlock be clearer about the fact that his book does not establish the superiority of the layman over the expert (Caplan 2007b).

Do Tetlock's results imply that there is no added value to expert advice compared to the judgment of well-informed laities? Concludes Tetlock:

In this age of academic hyperspecialization, there is no reason for supposing that contributors to top journals—distinguished political scientists, area study specialists, economists, and so on—are any better than journalists or attentive readers of the New York Times in “reading” emerging situations” (Tetlock 2005: 223).

In reply to this, Caplan argues that one should not misinterpret the meaning of Tetlock's results. According to him, all that Tetlock shows is that experts are bad at answering difficult questions, not easy ones, which does not imply that laypeople would do much better on either types of questions (Caplan 2007). Fair enough, but that still not does give us a decisive argument why we should ultimately trust economist experts more than laypeople (or their representatives) when it comes to making political decisions, including when those decisions have an economic component. In fact, the argument from diversity presented above implies that lack of cognitive diversity among experts can offset the advantage represented by their individual expertise, while, on the contrary, the cognitive diversity of large groups of non-experts can to a degree compensate for their lack of individual expertise. In terms of predictive accuracy, large groups of laypeople and small groups of experts may well draw a tie.

There are at least four different standards in the book, serving as benchmarks of citizens' biases: objective facts with a verifiable answer, the simulated “enlightened preferences” of a public with high political IQ, the simulated preferences of an “enlightened public” with the knowledge of a Ph.D. in economics, and finally the policy preferences of economists themselves. The problem is that objective facts are not a conclusive standard (the relationship between the possession of factual knowledge and

epistemic competence being too shaky), that taking economists' knowledge as the standard begs the question of who has authority in politics in the first place, and that the other two—"enlightened preferences" or "enlightened public"—are in fact slight variations about either facts or expert knowledge.

Conclusion

I have defended the view that as a collective decision-procedure democracy is more epistemically reliable than oligarchy. This is so because, even assuming that we could identify the best few, they would either not be numerous enough, and therefore cognitively diverse enough, to compete with many averagely smart people (in a direct democracy), or they would not be cognitively diverse enough over the long run (in a representative democracy).

I have proposed arguments supporting the hypothesis that deliberation and majority rule have epistemic properties of their own, which are maximized when their use is most inclusive, because of the key factor of cognitive diversity. Combining the epistemic properties of deliberation and majority rule, I conclude that democracy—in theory—is superior to any version of the rule of the few, including when we make unrealistic assumptions about the intelligence of the few. On that view, the good thing about democracy is that it naturally economizes on individual intelligence, while maximizing through sheer numbers the key factor of cognitive diversity.

Remember now that we neutralized the impact of two other factors of collective epistemic competence, namely virtue and information level, stacking the deck against democracy in the first case and, to some perhaps, for democracy in the second case.

Holding both the virtue and information factors constant, the rule of the many is epistemically superior to the rule of the few. What happens when we (theoretically) reintroduce those two variables?

If we reintroduce the virtue dimension all other things being equal, it should be obvious that it harms dictatorship and oligarchy more than it harms democracy. It would take saints in a dictatorship or an oligarchy not to abuse an unchallenged power to do what can best serve the ruler(s), even if they have to make some concessions to the masses to keep them quiet. By contrast, it is a long standing argument that the rule of the many, which is structurally designed to rule for the greatest number, economizes on virtue. To the extent that collective epistemic competence is also a function of the decision-makers' virtue, democracy is a fortiori preferable to the rule of the few when we reintroduce the virtue component.

If we now reintroduce the factor of (raw) information, what happens? On the one hand, democratic citizens have arguably fewer incentives to get informed than oligarchs, since their vote matters less to the outcome and since they only have to bear an infinitesimal cost for their decisions. Notice that this does not, however, necessarily apply to the decisions of representatives, whose votes can be pivotal and who are judged on their ability to deliver good results. For these reasons, representatives need to become informed and, conversely, to inform their constituencies. Further, it remains an open question whether low levels of information directly translate into low epistemic competence, especially if the relevant epistemic competence consists only in identifying competent representatives or answering general questions on referenda. The democratic mechanisms of deliberation and voting might be precisely why citizens need not become

more informed individually, if those mechanisms are able, as I hypothesize, to turn their relatively weak input into a much better output.

On the other hand, we saw that the great property of deliberation among many diversely thinking individuals is to process the available information more efficiently than deliberation among the few like-minded and thus produce additional, refined information. As long as voting occurs after sufficient public debates, one can argue that democracy is at least as well off in terms of the information available at the level of the group as an oligarchy would be. Whether the amount of information made collectively available through democratic deliberation more than compensates for voters' disincentives to become informed remains an open question. To what extent this problem of information really matters is not at all clear either. All in all, I do not think that reintroducing the information variable harms democracy or gives oligarchy an advantage.

The argument put forward in this paper forms a theoretical, autonomous argument in favor of democracy, distinct from arguments relying on theories of consent or equality or justice. This is not the place to defend the superiority of an epistemic justification of democracy over other justifications. Let me just suggest that whatever might be, or might have been, the initial reasons to prefer democracy over dictatorship or oligarchy, collective wisdom (democratic reason) might help explain why we keep it. Josiah Ober (this volume and 2009) argues on the basis of historical evidence that the superiority of Athens over rival city-states came from the epistemic properties of its democratic institutions, in particular the deliberative institution of the Council of 500 and the non-deliberative practice of "ostracism." I see his contribution as historical evidence supporting the theoretical claim presented here.

Let me add a final word on the conditions for democratic reason to emerge. I have insisted on the importance of cognitive diversity for the emergence of the phenomenon of collective intelligence. Without it, the mechanisms of deliberation and majority rule risk producing democratic unreason. I have assumed throughout this paper that more people bring in more cognitive diversity. In order for this correlation between numbers and cognitive diversity to remain plausible though, one must be considering a specific kind of society, characterized, among other things, by the existence of a free market of ideas, ensuring that the constant conflict of points of view and arguments renews perspectives, interpretations, heuristics, and predictive models—the toolbox of democratic reason. The emergence of democratic reason is thus conditional on the existence of a social and cultural context that nurtures and protects, among other differences, cognitive differences.

Although I cannot substantiate that claim here, my guess is that to the extent that the epistemic argument for democracy is true, it shows that democratic reason and liberalism go together. In other words, democracy is more likely to be smart if it is, also, liberal and applies to an “open society.” Illiberal or authoritarian democracies that foster conformism of views and stifle dissent risk turning both deliberation and majority rule into dangerous mechanisms for collective unreason, depriving themselves in particular of the possibility to come up with efficient solutions to collective problems, accurate information-aggregation, and reliable predictions. Other key factors are probably the independence of the media, as well as an educative system nurturing cognitive differences and the ability to express them.

Bibliography

Aristotle. *Rhetoric*.

Austen-Smith, David, and Jeffrey S. Banks. 1996. "Information Aggregation, Rationality and the Condorcet Jury Theorem," *The American Political Science Review* 90 (1): 34-45.

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.

Bohman, James. 2007. "Political Communication and the Epistemic Value of Diversity : Deliberation and Legitimation in Media Societies." *Communication Theory* 17 (4): 348-355.

Bovens, Luc and Wlodek Rabinowicz. 2006. "Democratic answers to complex questions—an epistemic perspective," *Synthese* 150: 131-153.

Breton, Philippe. 2006. *L'incompétence démocratique. La crise de la parole aux sources du malaise (dans la) politique*. Paris : La Découverte.

Caplan, Bryan. 2007. *The Myth of the Rational Voter. Why Democracies Choose Bad Policies*. Princeton and Oxford: Princeton University Press.

--2007. "Have the experts been weighed, measured, and found wanting?," *Critical Review* 19 (1): 81-91.

Condorcet, Marquis J. A. C. 1875. "Essai sur l'application de l'analyse à la probabilité rendue à la pluralité des voix," in *Les Archives de la Révolution Française*, Oxford: Pergamon Press.

Cohen, Joshua. 1989. "Deliberation and Democratic Legitimacy," in A. Hamlin and P. Pettit, eds., *The Good Polity* (Oxford: Blackwell): 17-34.

--1986. "An Epistemic Conception of Democracy." *Ethics* 97 (1): 26-38.

Condorcet, Marie Jean Antoine Nicolas de. 1976. *Selected Writings*. Ed. Baker, Michael K. Indianapolis: Bobbs-Merrill.

Converse, Philip E. 1990. "Popular representation and the distribution of information." In J.A. Ferejohn and J.H. Kuklinski (eds), *Information and Democratic Processes*, Chicago: University of Illinois Press.

- Cox, Gary. 2006. "The organization of democratic legislatures," in B. Weingast and D. Wittman (eds), *The Oxford Handbook of Political Science*, Oxford: Oxford University Press.
- Dekel, Eddie and Michele Piccione. 2000. "Sequential Voting Procedures in Symmetric Binary Elections," *Journal of Political Economy* 108 (1).
- Dowding, Keith and Martin Van Hees. 2007. "In Praise of Manipulation," *B. Journal of Political Science* 38: 1-15.
- Duxbury, Neil. 1999. *Random Justice: On Lotteries and Legal Decision-Making*. Oxford: Oxford University Press.
- Elster, Jon. 2008. *Le désintéressement. Critique de l'homme économique*. Paris: Le Seuil.
- 1989. *Solomonic Judgments: studies in the limits of rationality*. Cambridge, New York: Cambridge University Press; Paris: Editions de la Maison des sciences de l'homme.
- 1986. "The Market and the Forum : Three Varieties of Political Theory," in J. Elster and A. Aanund (eds), *The Foundations of Social Choice Theory*, Cambridge, Cambridge University Press.
- 1977. "Ulysses and the Sirens : A theory of imperfect rationality," *Social Science Information* 1977 (16) : 469-526.
- Elster, Jon and Hélène Landemore. 2008. "Ideology and Dystopia," *Critical Review* 20 (3): 273-289.
- Estlund, David. 2007. *Democratic Authority. A Philosophical Framework*. Princeton: Princeton University Press.
- 1997. "Beyond Fairness and Deliberation: The Epistemic Dimension of Democratic Authority." In James Bohman and William Rehg, eds, *Deliberative Democracy*. Cambridge, MA: MIT Press.
- 1994. "Opinion Leaders, Independence, and Condorcet's Jury Theorem," *Theory and Decision* 36 :131-162.
- Fauré, Christine. 1989. « La pensée probabiliste de Condorcet et le suffrage féminin. » In Crépel, Pierre, and Christian Gilain (sous la direction de), *Colloque international. Condorcet : mathématicien, économiste, philosophe, homme politique*. Paris, Minerve: 349-354.
- Galton, Francis. 1907. "Vox Populi," *Nature* 75 (March 7): 450-451.
- Goodwin, Barbara. 1992. *Justice By Lottery*. Chicago: The University of Chicago 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.
- Feddersen, Timothy, and Wolfgang Pesendorfer. 1998. "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," *American Political Science Review* 92 (1): 23-35.
- 1997. "Voting Behavior and Information Aggregation in Elections With Private Information," *Econometrica* 65 (5): 1029-58.
- Fishkin, James. 2009. *When the People Speak. Deliberative Democracy and Public Consultation.* Oxford: Oxford University Press.
- Fishkin, James and Bruce Ackerman. 2004. *Deliberation Day.* New Haven: Yale University Press.
- Goodwin, Barbara. 1992. *Justice By Lottery.* Chicago: The University of Chicago Press.
- Gottfredson, Linda S. 1997. "Why g matters: the complexity of everyday life." *Intelligence* 24 (1).
- Hong, Lu and Scott Page. 2012 (In Press). "The Micro-foundations of Collective Wisdom," in H. Landemore and J. Elster (eds), *Collective Wisdom: Principles and Mechanisms*, Cambridge, Cambridge University Press.
- 2009. "Interpreted and Generated Signals." *Journal of Economic Theory* 144 (5): 2174-2196.
- 2004. "Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers." *Proceedings of the National Academy of Sciences*, 101 (46): 16385-89.
- 2001. "Problem Solving by Heterogeneous Agents." *Journal of Economic Theory*, 97 (1): 123-63.
- Hutchins, Edwin. 1995. *Cognition in the Wild.* Cambridge, MA: MIT Press.
- Karotkin, Drora, and Jacob Paroush. 2003. "Optimum Committee Size: Quality-versus-Quantity Dilemma." *Social Choice and Welfare* 20 (3): 429-441.
- Kemp, Simon. 2007. "Psychology and opposition to free-trade," *World Trade Review*, 6: 1, 25-44.
- Kruger J. et D. Dunning 1999: 1121-34), "Unskilled and unaware of it", *Journal of Personality and Social Psychology* 77 (1999), 1121-34

- Ladha, Krishna. 1992. "The Condorcet Jury Theorem, Free Speech, and Correlated Votes." *American Journal of Political Science* 36 (3), 617-634.
- Lamberson, P.J. and Scott Page. Forthcoming. "Optimal Forecasting Groups," *Management Science*.
- Landemore, Hélène. 2009. "Majority Rule and the Wisdom of Crowds: The Task-Specificity of Majority Rule as a Predictive Tool," Paper Presented at the Paris International Conference on Majority Rule, Collège de France, May 13-15, 2009.
- 2008. "Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many" (Ph.D. Dissertation, Unpublished, Defended at Harvard University on December 13, 2007)
- Lave, Jean. 1988. *Cognition in Practice: Mind, Mathematics, and Culture in Everyday Life*. New York, NY: Cambridge University Press.
- List, Christian. 2004. "On the Significance of the Absolute Margin," *British Journal for the Philosophy of Science* 55 (3): 521-544.
- List, Christian and Philip Pettit. 2005a. "Group Agency and Supervenience," *Southern Journal of Philosophy*.
- 2005. "On the Many as One," *Philosophy and Public Affairs*, 33 (4): 377-390.
- 2002. "Aggregating Sets of Judgments: an Impossibility Result," *Economics and Philosophy* 18: 89-110
- List, Christian and Robert E. Goodin. 2001 "Epistemic Democracy: Generalizing the Condorcet Jury Theorem." *Journal of Political Philosophy* 9 (3), 227-306.
- 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.
- 2002. "Does Democratic Deliberation Change Minds?" prepared for the Workshop, Democratic Theory: The Canberra Papers, 27 March 2002.
- 2008. "Astroturing Infotopia." *Theoria*.
- Manin, Bernard. 1995. *Les Principes du Gouvernement Représentatif*. Paris : Flammarion.
- Mansbridge, Jane. 2009. "Deliberative and Non-Deliberative Negotiations" (April 6): HKS, Working Paper No. RWP09-010. Available at SSRN: <http://ssrn.com/abstract=1380433>

- 2003. "Rethinking Representation." *American Political Science Review* 97 (4): 515-528.
- 1999. "Should Blacks Represent Blacks and Women Represent Women? A Contingent 'Yes'." *The Journal of Politics* 61: 628-57.
- 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.
- Mulgan, Richard G. 1984. "Lot as a Democratic Device of Selection." *Review of Politics* 46: 539-560.
- Norman, D.A.1991. "Cognitive Artifacts." In Carroll JM (ed) *Designing Interaction: Psychology at the Human-Computer Interface*: 17-38. New York, NY: Cambridge University Press.
- Ober, Josiah. 2009. *Democracy and Knowledge*. Princeton and Oxford: University of Princeton Press.
- O'Leary, Kevin. 2006. *Saving Democracy: A Plan for Real Representation in America*. Stanford: Stanford University Press.
- Page, Benjamin and Marshall M. Bouton. 2006. *The Foreign Policy Disconnect. What Americans Want from Our Leaders but Don't Get*. Chicago and London: The University of Chicago Press.
- and Robert Y. Shapiro. 1992. *The Rational Public. Fifty Years of Trends in Americans' Policy Preferences*. Chicago, IL: the University of Chicago Press.
- Page, Scott. 2006. *The Difference. How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton: Princeton University Press.
- Panning, William H. 1986. "Information Pooling and Group Decisions in Nonexperimental Settings." In *Information Pooling and Group Decision Making*. Ed. Bernard Grofman and Guillermo Owen. Greenwich, CT: JAI Press.
- Pauly, Marc and Martin van Hees. 2006. "Logical Constraints on Judgment Aggregation," *Journal of Philosophical Logic* 35 (6): 569-585.
- Pettit, Philip. 2004. "Groups with a mind of their own" in Frederick Schmitt (ed.) *Socializing Metaphysics*, New York, Rowman and Littlefield: 167-93.
- Popkin, Samuel. 1994. *The Reasoning Voter*. Chicago: The University of Chicago Press.

- Popkin, Samuel, and Michael Dimock. 1999. "Political Knowledge and Citizen Competence." In Stephen Elkin and Karol Soltan (eds), *Citizen Competence and Democratic Institutions*. University Park: Pennsylvania State University Press.
- Rawls, John. 1993. *Political Liberalism*. New York: Columbia University Press.
- Rehfeld, Andrew. 2010. "On Quotas and Qualifications for Office," in Ian Shapiro, Susan Stokes, Elisabeth Woods, and Alexander Kirschner (eds), *Political Representation*, New York: Cambridge University Press.
- 2009. "Representation Rethought," *The American Political Science Review* 103 (2): 214-230.
- 2008. "Extremism in the Defense of Moderation: A Response to My Critics," *Polity* 40 (2): 254-271.
- 2006. "Towards a General Theory of Political Representation," *The Journal of Politics* 68 (1).
- 2005. *The Concept of Constituency: Political Representation, Democratic Legitimacy and Institutional Design*. Cambridge: Cambridge University Press.
- Rumelhart DE, P Smolensky, JL McClelland, and GE Hinton. 1986; "Schemata and sequential thought processes in PDP models." In: McClelland JL, Rumelhart DE, and the PDP Research Group (eds) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. II, "Psychological and Biological Models," pp. 7-57. Cambridge MA: MIT Press.
- Sabbagh, Daniel. 2003. *L'égalité par le droit. Les paradoxes de la discrimination positive aux Etats-Unis*. Paris : Economica.
- Schofield, Norman. "Madison and the Founding of the Two-Party System." in *James Madison: The Theory and Practice of Republican Government*, ed. Samuel Kernell. Stanford, CA: Stanford University Press 302, (2002): 27.
- Sintomer, Yves. 2007. *Le pouvoir au peuple: jury citoyens, tirage au sort et démocratie participative*. Paris: La Découverte.
- 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.
- Stone, Peter. 2010. *The Luck of the Draw: The Role of Lotteries in Decision-Making*. Oxford: Oxford University Press.

--2009. "The Logic of Random Selection." *Political Theory* 37: 375-97.

--2007. "Why Lotteries Are Just?" *The Journal of Political Philosophy* 15 (3): 276-295.

Urbinati, Nadia. 2006. *Representative Democracy: Principles and Genealogy*. Chicago: University of Chicago Press.

Wertsch, J.V. 1998. *Mind as Action*. New York, NY: Oxford University Press.

Young, H.P. 1988. "Condorcet's Theory of Voting," *American Political Science Review* 82: 1231-44