

Democratic Reason: the Mechanisms of Collective Intelligence in Politics

Hélène LANDEMORE

Department of Government
Harvard University
Cambridge, Massachusetts

A.T.E.R.
Collège de France
Chaire de M. Elster
Paris, France

helene.landemore@college-de-france.fr

Prepared for Presentation at the Conference “Collective Wisdom: Principles and Mechanisms”,
Collège de France, Paris, May 22-23, 2008

Democratic Reason: the Mechanisms of Collective Intelligence in Politics
Hélène Landemore

Abstract: This paper argues that democratic institutions can be seen as ways to channel “democratic reason,” or the distributed intelligence of the many—a concept that I build in part on the insights of the cognitive sciences. I argue 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 picks the better political answer within a given a context and a set of values. Under the conditions of a liberal society, characterized among other things by sufficient cognitive diversity, these two mechanisms give the rule of the many an epistemic edge over any variant of the rule of the few.

Introduction

The idea of “democratic reason,” roughly the collective intelligence of the people, put forward in this paper might seem immediately problematic, some might even say contradictory. If we take seriously the idea of collective wisdom advanced in this colloquium, however, we need to consider that the intelligence of the many might be one of the reasons why democracy works and why democracy is valuable. In this paper I propose a follow up on the work of David Estlund, who has in my view definitely carved out the space for knowledge-based arguments for democracy, and I propose what will no doubt be considered a bold, perhaps even overly optimistic claim. David Estlund (2007) claims that democracy epistemically beats a random decision-procedure. I am interested in a more ambitious comparison, that with the real political alternatives: rule of one and rule of the few.¹ I make the claim that given conditions of sufficient education and freedom amongst the members of a regular group of human beings, the rule of the many beats any alternative of the rule of the few as a decision-procedure regarding collective choices for that group. I generalize this claim to argue that the rule of the many epistemically beats any variant of the rule of the few. The boldness of the

¹ This comparison is precluded in Estlund’s view by the problem of invidious comparisons (between people’s claim to know better) and his general premise that political entitlement does not simply come from the fact that one knows best. I disregard this premise here, which allows me to pursue the larger comparison. I think that comparisons are invidious but not impossible and that they lead to believe that the group as a whole, rather than any subset within it, knows better. For a defense of the idea that political authority might well be if not epistemic through and through, at least essentially so (that is essentially based on an ability to know better, rather than on consent), see chapter Three of my doctoral work “Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many” (Harvard University 2007).

claim comes in part from the fact that I interpret oligarchy and dictatorship in the best possible light, granting that the oligarchs and the dictator pursue the good of the greatest number as opposed to their narrow interests.

While the comparison between the rule of many and the rule of one should be fairly intuitive (many heads are better than one), the second claim is more controversial. If collective intelligence 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 beat the rule of the average citizen. In this paper I claim that this is not the case because of the importance of cognitive diversity—roughly the ability to see the world from different points of view—in the emergence of collective intelligence. Applying the theoretical findings of Scott Page (2006) about the relative importance of cognitive diversity and individual ability for collective problem-solving and predictions, I argue that since an oligarchy of even very smart rulers will always display less cognitive diversity than a direct or representative democracy, the few can in general not match the epistemic competence of many moderately smart but diversely thinking individuals. The power of the argument presented here is thus that even granting the possibility of identifying the smartest individuals in a given population and even granting their virtue, I propose that democracy would still be preferable to oligarchy because the few best would not be numerous enough, and therefore cognitively diverse enough, to compete with many averagely smart people.

In order to support that theoretical claim, I study two of the decision-mechanisms that give democracy an epistemic edge: maximally inclusive deliberation and majority rule under universal suffrage. I argue that, to the extent that cognitive diversity is correlated with the number of participants, deliberation among the many epistemically beats deliberation among the few. I also argue that given the same assumption of cognitive diversity brought by

numbers, aggregation of judgments among the many epistemically matches 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 procedures give democracy an epistemic edge over any variant of the rule of the few.

This paper deliberately avoids considering potential epistemic failures in deliberative settings and voting, such as the well-documented problems of group think or information cascades, or the problem of tyrannical majorities. To the extent that these problems exist and would arise whether the rulers are few or many, they are not discriminating in the comparison attempted here. Further, epistemic properties are only probabilistic and nowhere do I argue here that rule of the many (or one or a few) is infallible. The question of the proper institutional answers to bring to those epistemic failures and risks, however important, cannot be addressed here.

This paper has four sections. In the first section, I propose a definition of the few key-concepts of my general claim about democratic reason. 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. I consider both the case of direct and representative democracy. The third section turns to majority rule. I consider successively three plausible theoretical accounts of its epistemic properties: the Condorcet Jury Theorem, the Miracle of Aggregation, and the law “The crowd beats the average” (Page 2006). I also deal in this section with an objection related to the fact that in a democracy voters have no incentives to become informed, which arguably considerably lowers their epistemic competence (Caplan 2007). The fourth section recapitulates the epistemic advantages of the rule of the many over versions of the rule of few, including when one reintroduces the virtue and information factors.

1. *Democratic reason and epistemic competence.*

I define democratic reason as a certain kind of distributed collective intelligence specific to democratic politics. This concept is meant in part to contrast with the less inclusive (in my view) 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. I also develop the concept of democratic reason in relation to the idea of “collective wisdom” or, as I will prefer to call it, “collective intelligence,” building here on a few insights from psychology and the cognitive sciences. The first lesson I derive from the cognitive sciences is that intelligence is a broad ability for comprehending our surroundings and figuring out what to do, as opposed to mere book learning or test-taking smarts (Gottfredson 1997: 13). It is thus irreducible to IQ or other single-dimensional feature. Collective intelligence is such a complex notion of intelligence premised of groups as opposed to individuals.

The second insight is that although collective intelligence can theoretically be a direct function of individual intelligence (the sum of the parts), it is more aptly seen as an “emergent” or “supervenient” property (more than the sum of the parts). Examples of supervenient intelligence are displayed by groups of social animals, such as ants or bees, which display a form of intelligence at the level of the group that is not be found at the level of each distinct animal. Applied to humans, the idea of collective intelligence does not necessarily imply that no intelligence is to found at the level of individuals, but that the emerging product is nonetheless of a different nature or quality.² As we will see, an essential

² I specify that this idea of collective intelligence does not commit me to a form of methodological or philosophical holism.

ingredient of collective intelligence is the cognitive diversity of the group, a property which by definition cannot be found in individuals.³

The third lesson from the cognitive sciences that I apply is that cognition and intelligence can be “distributed” (Lave 1981) over both space and time. Spatially speaking intelligence can be distributed over individuals and what is called “cognitive artifacts,” i.e. devices that help us accomplish complicated tasks and/or make smarter decisions or choices.⁴ A famous study by the cognitive anthropologist Edwin Hutchins (1995) shows for example how the computation involved in steering a large ship does not take place in the head of any particular individual, not even the captain but in the coordination of the minds of different individuals equipped with navigational artifacts, such as landmarks, maps, phone circuits, and organizational roles.⁵ The relevant cognitive unit in this case is no particular individual, but the system “crew+relevant cognitive artifacts” as a whole.

Some cognitive artifacts also allow cognitive processes to be broken down into successive tasks, distributing the cognitive effort not just through space but through time as well. For example, a list of things to do divides the cognitive process of remembering into at least three chronological steps: the construction of the list, the mental action of remembering to consult the list, and finally the actual reading and interpretation of the items on the list (Norman 1991: 21).

Applying the related notions of distributed collective intelligence and cognitive artefact to collective decision-making, I propose to call “democratic reason” the kind of

³ Although of course some analog of cognitive diversity within a group can exist within individuals when they make a conscious effort to see the world from different viewpoints, but that is a separate point.

⁴ According to the exact definition, cognitive artifacts are “those artificial devices that maintain, display, or operate upon information in order to serve a representational function and that affect human cognitive performance” (Norman 1991: 17). Examples are language, inscriptional systems for representing language, maps, lists and calculators. The great advantage of cognitive artifact is that all that the individual has to know in order to solve a problem is how to use the cognitive artifact, which contains in itself the knowledge required to solve this particular problem. Thus to solve a complex multiplication, all I need to know is how to spatially display the digits on a sheet of paper and perform simpler calculus in a given sequence (Rumelhart et al., 1986 and Wertsch 1998). I can also use the superior cognitive artifact that is a calculator.

⁵ This example is not meant as an example of democratic organization, just as an example of collective distributed intelligence (but distributed collective intelligence need not be democratic).

distributed collective intelligence specific to democratic politics. 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 this approach is that it allows us to explain how the individual citizen cognitively unburdens him or 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 (in which case, in my view, the “right” of the people to self-rule finds itself 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 (and probably other political cognitive artifacts that I am not considering).⁶ Another way to say this is that democratic decision-procedures are societal ways of making collective decisions that economize on the epistemic competence of the individuals taking part in the decision.

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, a dictator as well), I will need to explain why the combination of democratic deliberation and majority rule with universal suffrage beats the combination of deliberation among the few and majority rule among the few.

⁶ 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. This will prove particularly useful if we believe that democracies should be able to learn, particularly from their own mistakes, how to immunize themselves against the worst forms of individual and cognitive failures.

Let us for now turn briefly to the notion of “epistemic competence.” Epistemic competence is the competence one has in virtue of knowing something or having a certain kind of knowledge. It must first be distinguished from another form of competence, which one might call “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. Notice that I (deliberately) stack the deck in favor of oligarchy and dictatorship here.

Epistemic competence is also distinct from something with which it is often confused, namely 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 too often reduced to a simple equation by political scientists, as if the level of information was all that mattered to predict epistemic competence. In fact 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 incompetent. Studies about the “reasoning voter” and her use of heuristics and cognitive shortcuts to vote based on little information what she would have voted based on more exhaustive information (e.g., Popkin 1994) have complicated the picture but this identification political competence and level of information remains nonetheless the general background assumption of many approaches to political competence. I will assume, somewhat controversially but following here the most recent findings in cognitive sciences and some authors in political sciences (e.g. Lupia 2006), that epistemic competence is so minimally correlated with the level of

information as it is measured by existing empirical surveys that I can neutralize this factor. I will instead 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. One might object that by doing I stack the deck in favor of democracy. I deny that this is the case but postpone a detailed answer to that objection to the third 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. 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 cognitive artifacts processing this information for the group. I argue that democracy does not reduce the collective level of information (to the contrary) and that it is dubious that democratic citizens on average individually fall below the threshold that allows them to vote competently anyway.

Finally, one must distinguish between collective epistemic competence and individual epistemic competence. As I said earlier, collective epistemic competence might be more than just the sum of individual epistemic competence, and indeed a property emerging from the right mix of individual epistemic competence and some other things. 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. To be clear:

$$(1) \text{ Collective Epistemic Competence} = f(\text{individual epistemic competence, cognitive diversity of the group})^7$$

⁷ I assume that the function f reflects the right degree of interaction between individuals. Since I assume oligarchy and dictatorship to be the best they can, I can afford to assume that the type of interactions between democratic citizens is of the kind that allow for individual epistemic competence and the cognitive diversity of the group to matter.

Once we have an answer to the question of which decision-procedure has the higher collective epistemic competence given equal amount of information and virtue in the rulers, it will be easier to reintroduce the “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 makes democracy an epistemically competent decision-rule is inclusive deliberation (that is deliberation that involves, directly or indirectly, all the members of the group). Before I turn to why, for a given group, democratic deliberation beats deliberation among the few, I need to define deliberation and establish its general epistemic properties.

Deliberation, according to Aristotle, consists in an exchange of arguments for or against something (Aristotle, *Rhetoric*, I, 2). Contemporary deliberative democrats have added to that definition the goal of a rational agreement or consensus on the better answer or argument. Deliberation is also opposed to voting or bargaining and 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 obvious epistemic superiority.⁸

Three main arguments have been classically advanced for the epistemic properties of deliberation. 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.

⁸ I summarize here a huge literature on the subject, following a recent survey by Martí 2006.

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 other is an adaptation of an example given by Scott Page.

In the film (by Sidney Lumet, based on the play by Reginald Rose) one brave dissenting jury member—number 8, played by the actor Henri Fonda—manages to persuade the other 11 jurors to reconsider the guilty sentence 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 violent slum, is the one who notices that the suspect could not possibly have stabbed his victim with a switch-blade. No other juror was acquainted with the proper way to use a switch-blade. Juror number 9, who is an old man, then questions the plausibility of the time it took one of the key witness to cross the corridor. He too contributes to changing the collective perspective on the way the crime took place. One of the most rational jurors, a stock broker who is not convinced by any of the other arguments, finally has to admit that a shortsighted 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 participant contributes 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 witness—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 short-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.⁹ 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

⁹ I borrow and suitably modify an example from Scott Page (personal communication).

following respective local optima, 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))

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, his only optimum, which has a value of 9. 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 idea, but then Paris (8) is better.” The député from Calvados has to agree (forceless force of the better argument oblige). 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 three others (Marseille, Paris and Caen). 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 and Corrèze would not necessarily have thought of Caen (10) if they had not been spurred to do so by the député of Pas de Calais who proposed Grenoble (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 Scott Page’s “Diversity trumps Ability Theorem,” what matters most to collective intelligence in 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 portioning perspectives), diversity of heuristics (the way of generating solutions to problems), and diversity of predictive models (the way of inferring cause and effect) (Page 2006: 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 2006: 163). (The four conditions are fairly reasonable—I refer the reader to Scott Page for the details¹⁰).

The general point 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 “non-guilty” reached by the group. Similarly if all three députés were thinking exactly alike—say like the député of Pas de Calais—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).¹¹ 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.

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

¹⁰ 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 is that any solution other than the global optimum is not a local optimum for some nonzero percentage of problem solvers. This 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 2006: 159-162.

¹¹ In fact, given the way I defined the probabilities, having more members would not make us smarter in any way since we would have the exact same probability of reaching 7 or 9 as a group as we would have as distinct individuals. Notice, however, that if we are both more likely to get to 9 than 7, our group is more likely to reach 9 than any of us individually (due to the law of large numbers)—I shall turn to the demonstration of that point on the section on majority rule.

examples I gave occur among less than 11 people. 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?

I will first argue that to the extent that increasing the numbers of participants in a group ensures greater cognitive diversity, more is smarter.¹² I thus hypothesize that one can 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 are even more cognitively diverse, and thus smarter, than three. Similarly, if 11 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 seems more intuitive than the reverse assumption.

The next problem, however, which might dampen our enthusiasm for numbers, is that deliberation involving all those concerned is not always feasible. In effect, passed 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. Luckily the options here are not restricted to oligarchy, since the device of representation allows the indirect or mediated involvement of the many in a decision taken by the few. An assembly of representatives is indeed distinct from a group of oligarchs thanks to two institutional principles: periodic elections and accountability (Manin 1996).

¹² 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).

Representatives are elected to the position of decision or law-makers, as opposed to born into it (like aristocrats) or appointed by one or a few persons only (as might be the case of non elected magistrates or experts). Elections, it might be objected, do in practice retain an oligarchic, even aristocratic, flavour in that they involve a principle of selection, often on criteria that give more chances to the more educated and/or the richest members of society, who then tend to stay in power and self-reproduce. As a consequence, the individual ability of those groups might be high but it is unlikely that their cognitive diversity will be too. Some authors have suggested selection by random lottery as an alternative to election (e.g., Elster 1989: 78-103) and argued that lotteries are actually in principle more democratic and representative (e.g., Mulgan 1984: 539-560, Goodwin 1992, Duxbury 1999, Stone 2007 and Sintomer 2007). Not only are lotteries arguably more just, 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. On the other hand, there are problems of incentives, motivations, and accountability (since selected people's hope to be re-selected is independent from what they achieve while in power) that render this idea difficult to implement. Election might be just the second-best in terms of reproducing cognitive diversity. The fact that elections are periodic (every four or five years) also ensures a minimal renewal of the rulers over time (although lesser than would be the case with lotteries), which is crucial in preventing representatives from turning into a group of similarly thinking oligarchs.

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 throughout. This claim is both normative and empirical. It is normative, because on a democratic (as opposed to a Burkean) reading of their role (e.g., Urbinati 2006), representatives are supposed to take into account their constituents' interests and judgments, not act and decide entirely on their own. It is a

practical reality, because representatives' re-election being conditional on constituents' satisfaction with what they are perceived as having accomplished while in power, these representatives have an incentive to constantly keep in touch with public opinion.

Representatives are not enslaved to their constituents. They can use their judgments as to the best means to the given results on which their constituents will judge them; they also have some wiggle room to the extent that failure to deliver pre-election promises can always be blamed on some other factors than their own incompetence. Nonetheless, representatives generally pay careful attention to polls and whatever signs—strikes, demonstrations, and even riots¹³—translate public opinion. I thus propose a non-orthodox reading of representation as the institutional trick that allows for democratic deliberation on a human scale while preserving at least some of the cognitive diversity present in the entire group. I do not make the optimistic (Burkean) hypothesis that representation ensures rule by an aristocracy of merits and talents.¹⁴

The conclusion of what precedes should be clear: deliberation involving the many, in a direct form (where feasible) or an indirect form (through representation) 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.

3. Majority Rule under universal suffrage

¹³ Historically fairly normal elements of the political vocabulary in France for example.

¹⁴ 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 I think a more accurate and modest account of the function that representatives are fulfilling is 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.

Deliberation is far from being a perfect or complete decision-mechanism, in part because it is time-consuming and rarely produces unanimity.¹⁵ In most cases, it needs to be supplemented, if not replaced, 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 will argue that far from just being a fair way to settle disagreement about the choice of an option, majority rule is also a 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 beats majority rule used by a minority within the larger group.

I will consider three 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.” The latter is the most convincing explanation of why collective predictions are better than individual predictions. It also explains why majority rule is epistemically inferior to deliberation. Unlike deliberation, sheer aggregation of judgment does not allow for the weeding out of bad information and ideas from good ones. Majority rule averages out all expressed judgments. Consequently, all things equal otherwise, the

¹⁵ Deliberation is also imperfect in that it is subject to “group think,” polarization, informational cascades, or social pressure, but to the extent that these problems affect deliberation in general, I do not think they are of relevance for the comparison between inclusive and non inclusive deliberation once we admit that the feasibility problems introduced by large numbers is solved by the institutional device of representation. I deal at greater length with the epistemic failures and limits of both deliberation and majority rule in Chapter Five of my dissertation.

aggregated prediction of a large random group will only rarely predict more accurately than the aggregated prediction of a small group of experts.

3.1 The Condorcet Jury Theorem

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 (the enlightenment assumption); 2) they vote independently of each other (the independence assumption); and 3) they vote sincerely or truthfully (the sincere voting assumption). To briefly illustrate the power of large numbers harnessed by majority rule according to the CJT, consider 10 voters, each of which has a .51 probability to be correct on any yes or no question. A majority of 6 will have 52% chance of being right. Expand now the group to 1000 people and a majority of 600 hundred is almost 100% sure to be right. This is merely an implication of what is also known as the central limit theorem.

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. These analyses, however, usually fall short of drawing substantive normative implications, in part because the CJT is generally considered of little relevance to the real world. Its assumptions are analyzed as pure mathematical technicalities with no matching real-life application (e.g., Lhada 1992).¹⁶ Meanwhile, more philosophically oriented theorists remain unsure of the usefulness of Condorcet theorem for democratic theory. Even David Estlund, an unambiguous advocate of an (at least partially) epistemic theory of democracy, who repeatedly mentions the CJT in support of the superiority of majoritarian decisions over alternative decision-rules,

¹⁶ Duncan Black set the mood for this reception of the theorem in this respect. Despite a sincere admiration for Condorcet’s finding and its value for group-decision theory in particular, Black was skeptical about the wider applicability of the theorem (Black 1958: 185). He was particularly worried about the idea of a “probability of the correctness of a voter’s opinion”—a phrase that he considered “without definite meaning” (Black 1958: 163).

ends up distancing himself from the theorem as “less than trustworthy” (Estlund 1997, 189) and more recently as entirely “irrelevant” (2007 : chapter 12). In the following I try to establish the normative relevance of the theorem, by defending the empirical plausibility of its main assumptions: the enlightenment assumption, the independence assumption, and the sincerity assumption.

The “Enlightenment” Assumption or, Why Trust the People?

The first assumption assumes that the average voter can be trusted to be right more than half the time on the political issue at hand. How plausible is this hypothesis? Regardless of Condorcet’s own position,¹⁷ one might argue that this assumption is immediately plausible for at least a limited range of issues, such as choosing competent representatives or deciding on general policy orientations, such as those at stakes in referendums. This domain of competence includes what Lupia (2006) considers as “big choices.” “Big choices” have the following three properties : 1) they are generally binary (yes or no, or a choice between two candidates), 2) simple and widely available pieces of information (or cues) help voters vote as they would have if they knew more about the choices in question, and 3) their salience makes it easier for ordinary people to use cues efficiently. Given these three properties, Lupia concludes, big choices stack the deck in favor of competent voting, even by people who

¹⁷ Given the “prejudices” and “ignorance” common among people, including about “important matters,” Condorcet remarked:

[...] it is clear that it can be dangerous to give a democratic constitution to an unenlightened people. A pure democracy, indeed, would only be appropriate to a people much more enlightened, much freer from prejudices than any of those known to history (*Selected Writings*, 49).

Over the years, however, Condorcet came to endorse a more optimistic trust in the reasonability of the people and, consequently, universal suffrage. One reason for his change of mind is perhaps that he found empirical evidence for people’s gradual enlightenment during his own lifetime. The historian Michelet suggested that the 14th of July—the Bastille takeover—is the event that converted Condorcet to universal suffrage. Baker, more circumspect about the hypothesis of an overnight conversion, notices that Condorcet had a gradual evolution with respect to voting rights. He describes Condorcet moving progressively from the theoretically hardly consistent but pragmatic defense of limited suffrage, to a perfectly consistent but almost utopian (in most of his contemporaries’ view) defense of universal citizenship including the poor, women, and even black people. At least in the case of women, however, it is unclear that Condorcet was willing to grant them specifically the right to vote, as opposed to a more vague “right of citizenship.” For an excellent analysis of Condorcet’s ambiguities on that topic, see Christine Fauré 1989.

cannot answer common survey questions about politics. Lupia further suggests that it is probable that ordinary people are competent in a wider range of circumstances than many critics presume.

Contrary to the elitist reading of representative democracy, which assumes that the only competence ordinary citizens have is in identifying more competent elites that will rule on their behalf, I would further hypothesize that citizens could not identify in others a competence that they would be entirely deprived of. This hypothesis is supported by empirical findings in psychology demonstrating the double bind in which incompetent people often find themselves (J. Kruger et D. Dunning 1999: 1121-34). Not only are incompetent people in this study lacking a given skill or knowledge, but they are 1. unlikely to recognize their own incompetence (they tend to overestimate themselves), 2. unlikely to be able to identify other people's actual competence, 3. less likely than competent peers to be able to appreciate their actual level of competence from a comparison of their own choices and performances with those of others (as they cannot seem to make good use of available information).

The experiments just mentioned were run for the case of grammatical knowledge and humor, but I see no reason to think that it would be any different for political competence. Of course, this claim would need more empirical support to be established but I would then point out that the symmetric assumption that incompetent citizens can solve the problem of their own incompetence by electing competent people is just as empirically unfounded and much more implausible. I suspect that in a representative democracy we simply have the rulers that we deserve, namely people that are neither more nor less intelligent than the average citizen (except perhaps that they develop a certain expertise over time but that is a separate question since anyone put in that position would develop the same expertise too).

Even though I just suggested that incompetent people are most likely unable to identify competent people, it is more plausible that moderately competent people can identify

much more competent people. In other words, identifying competence might not require the exact same level of competence (but perhaps just a minimal amount of it). If so, what matters in a representative democracy is that regular citizens are smarter than a flip coin when it comes to choosing between candidates, and then that the selected representatives be smarter than a coin flip for all the relevant political questions that they have to deal with. Here a frequent assumption is that the selection of individuals through periodic elections is what ensures the epistemic superiority of representatives over regular citizens.

For the sake of the comparison attempted here, however, I will assume the worst case scenario for democracy, namely that representatives are not smarter than average citizens but that they simply form a category of people with the time and the energy required for political activities. The reason to delegate decisions to a smaller number of representatives are thus in my view not primarily epistemic, but practical: they have to do with the necessary division of labor in large, modern countries and the upper limit on the size of a functional deliberative assembly.

Some authors have suggested that there are domains where laymen and their representatives are objectively biased or incompetent (economic topics in particular, see Caplan 2006 and for an attempted psychological explanation of some amount of that incompetence see Kemp 2007, who analyzes the reasons behind regular people's fear of free trade). Assuming that this topical incompetence is true and can be demonstrated, one would still need to prove that the problem cannot be solved (if only over time) through education and pre-voting debates. If the incurability of economic incompetence could be established, then the conclusion would be, indeed, that rule of the few beats rule of the many for that particular domain. That does not suffice, however, to demonstrate the superiority of oligarchy over democracy more generally. It might just be that some questions are just too technical to be called "political." Whether or not economic questions belong to such a category is an open

debate. I would argue that for most political questions, including many economic ones, the average citizen has more common sense than a coin flip and so do the representatives chosen by them. On that (certainly optimistic) view, trust in the public does not require a leap of faith, but a simple belief in what Descartes deemed the most fairly distributed thing in the world.¹⁸

Independence

According to a commentator, the assumption of independent voting forms “the main weakness of the CJT,” because it seems to require that there be no opinion leaders, no communication among voters, not even the sharing of common information, culture, religion, or beliefs, or other elements that could lead to correlated votes (Lhada 1992, 621). On this very demanding reading of independence as lack of correlation between the votes, the Condorcet theorem becomes wildly implausible. Votes are likely to be correlated because in real life voters share common information, communicate with each other, and are influenced by various schools of thought or opinion leaders espousing the same or opposite opinions. Empirical studies (e.g. Panning, 1986) show that individual judgments are seldom independent because of mutual influence and communication.

Luckily the same commentator offers a generalization of the theorem to correlated votes: as long as the average interdependence between votes is sufficiently low, the theorem holds (Lhada 1992, 626). Relaxing the independence assumption into low vote correlation permits to envisage pre-voting communication and deliberation. There is thus no need to think that the Condorcet Jury Theorem requires endorsing the atomism of voter deliberation (everyone making up their minds alone) that some commentators have attributed to Condorcet on the basis of a strict reading of the independence assumption (Grofman and Feld, 1988;

¹⁸ Descartes’ assumption was not backed up by sophisticated empirical surveys and measurements but it has undeniably an empirical flavour, being based on the sample of people Descartes himself had encountered in his life.

Grofman, Owen and Feld 1989). A better way to read Condorcet's independence assumption is perhaps as an idealization of a low vote correlation made possible in a free and plural society (where a diversity of views are encouraged). On that view, independence translates as the flip side of the assumption of "enlightenment"—not the cognitive side tied to education, but the agentic side tied to autonomy, or the ability to make up one's minds independently, even if one cannot avoid reading the same newspapers as others.¹⁹

Sincere or truthful voting

The problem represented by the assumption of truthful voting is not neutralized by the assumption made in this paper that all rulers are pursuing the common good. Indeed rulers might sincerely pursue the common good, in non sincere ways. What is excluded, however, is self-interested voting, which is distinct from strategic and more generally non sincere voting. Let me explain these distinctions in turn.

Sincere voting is first distinct from self-interested voting. In the theoretical frame of the CJT (which is also that of this paper and more generally any epistemic approach to democracy (e.g. Cohen 1986)), voting is a judgment about the common good, not the expression of a self-interested preference or even a judgment about what is most likely to best serve *me*. A self-interested vote might be sincere, but it would miss the point. For a long time, economists and rational choice theorists in general have taken the assumption of self-serving voting as granted and the assumption of altruistic voting as utopian. The literature on voter's psychology, however, suggests that altruistic voting is more plausible than self-interested, pocket-book voting (e.g., Popkin 1994: 21). In fact, in light of overwhelming empirical evidence in favour of non self-interested voting, rational choice theorists are now arguing that it is rational, on the "right" understanding of economic rationality, to vote non-

¹⁹ Even so, one might argue that the assumption of independence assumes too much diversity across the signals that people get from the world and base their judgment on. I will return to that question shortly when discussing objections to the miracle of aggregation.

selfishly, that is altruistically and/or ideologically (for a recent defense of that idea, see Caplan 2007; for a critique of the axiom of self-interested behavior in economic theory more generally see Elster 2008 forthcoming).²⁰ If my vote makes little difference to the collective outcome and if I am unlikely to pay the price of my mistakes, I can indulge in ideological or expressive voting, feeling good at hardly any cost for myself.²¹

Sincere voting is more interestingly distinguished from different forms of non sincere voting, including strategic voting. In sincere or truthful voting, voters do not seek to hide their genuine preferences, or manipulate their votes either to get their interest (as in pocket-book voting) or that of someone else served (as in altruistic voting), or to merely throw off the result (as in what could be called “anti-social” voting). Some authors have recently argued that manipulation and strategic voting do not necessarily form objections to voting procedures, suggesting that these games of second-guessing and strategizing may actually contribute to the quality of the outcome (Dowding and Van Hees 2007). Let us however consider, as the framework of the CJT invites us to do, that we need sincere voting. How plausible is this assumption?

According to social choice theorists, not really: truthful voting does not make sense from a rational choice point of view (e.g. Austen-Smith and Banks 1996 and Feddersen and Pesendorfer 1997).²² This is so because, even in a group sharing a notion of the common good, “sincere voting does not constitute a Nash equilibrium” (Austen-Smith 1996, 34). Worse still, in the jury context, the probability of attaining the truth is actually maximized

²⁰ Jon Elster shows that the hypothesis of self-interest—applied to voting or other behaviors—is often psychologically implausible and in fact not even essential to economic theory.

²¹ This problem of incentives raises another kind of problem, namely that of lowering voters’ probability of choosing the right answer. In my view, this re-interpretation of voting as rationally ideological does not make it less interested or selfish, so that if we take seriously the idea of altruism, altruist voting should also be minimally smart. I will go back to this problem when I deal with Caplan’s critique of the miracle of aggregation.

²² For Austen-Smith and Banks, the assumption of sincere voting “is inconsistent with a game-theoretic view of collective behavior. A satisfactory rational choice foundation for the claim that majorities invariably “do better” than individuals, therefore, has yet to be derived.” (Austen-Smith and Banks 1997: 35).

when voters vote strategically (e.g., Dekel and Piccione 2000). The pursuit of truth at the level of the group and strategic voting are not contradictory in a situation of common goal.²³

Against this objection, one might point out that strategic voting is problematic in a context where strategic interactions matter, such as a small-sized jury (11 people). However, the interesting application of the jury theorem for democracy is in the case where large numbers are involved and strategic interactions disappear. I do not know where the threshold is exactly to be found, but I would venture that if there are more than 100 people involved, it becomes hard to keep track of who is voting what (assuming that this information itself is available) and strategize about the best use of one's vote in order for it to affect meaningfully the outcome. In such cases—which includes voting within an assembly of representatives as well as referenda—the assumption of truthful voting is at least theoretically supported. In mass elections, since my vote does not make a difference anyway, it is not irrational to vote truthfully (Austen-Smith and Bank 1996 and Feddersen and Pesendorfer 1997). It seems that, at the very least, the CJT's assumption of truthful voting for the cases of interest for us has yet to be disproved.²⁴ In fact, majority rule used with large numbers seems more immune to manipulation than majority rule used with the small number.

The assumptions of the CJT are not all equally plausible, the most difficult to justify being probably the independence assumption. If and when they are verified, however, the CJT supports the claim that many average people can be at least as smart as a few smart people. On that account majority rule used by an oligarchy and majority rule used in a democracy

²³ In relation to that question of strategic versus sincere voting, let me add a point on the possibility of agenda manipulation. The CJT assumes that choices are binary, which is sometimes regarded as implausible. In fact it is not implausible to the extent that an election involving several choices, say three candidates A, B, and C, can always be re-described as a sequence of binary choices (opposing for example A to B, then whoever wins to C; or A to C, and whoever wins to B; or B to C, and whoever wins to A). The problem of reducing a multiple-choice situation to a sequence of binary choice is the possibility of agenda manipulation. I do not know of a good answer to that problem of agenda manipulation.

²⁴ I am not aware of empirical works documenting the ratio of people who vote strategically from those who vote sincerely. I would assume the proportions vary with elections. I am not sure either what proportion of sincere to non sincere voting is sustainable for the CJT to still work but I assume that such a ratio exists and that is not too low.

(whether by the citizens or their representatives) draw a tie. The advantage of majority rule with universal suffrage is that it is less taxing in terms of individual intelligence.

3.2 *The “Miracle of Aggregation”*

The “miracle of aggregation” (Converse 1990; Caplan 2007) is another explanation for collective intelligence distinct from the Jury Theorem, although also involving the law of large numbers. The “miracle of aggregation” is an appealing account of why the average guess of large groups of people on matters with a factual answer tends to be uncannily accurate, as in the case of information-markets or Galton’s famous experiment at a weight-judging competition. To rehearse briefly the latter, now worn anecdote, the 19th century statistician was attending a country fair, in which one of the attractions was a game of guessing. The goal was to guess the weight of an ox once slaughtered and dressed. Galton took the answers of the 800 participants or so and computed the mean, which turned out to fall within one pound of the right answer.²⁵ Many other anecdotes, recounted in both Surowiecki (2004) and Sunstein (2006), vividly illustrate the same “miracle” of group intelligence.

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.²⁶ 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 one person in the crowd knows the right answer and all the others make mistakes that cancel each other out, the right answer is still going to rise to the surface, so to speak. Applied to the

²⁵ 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.

²⁶ This “elitist” version probably goes back to Berelson, Lazarsfeld, and McPhee (1954).

experiment of Galton, this explanation would imply that one or a few persons in the crowd knew the right answer and all the others made mistakes that cancelled each other.²⁷

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 weight, although none of them knew it exactly. Page and Shapiro apply this model to account for the rationality of public opinion.²⁸

A third version of the miracle of aggregation consists in seeing the right answer 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

²⁷ In fact, this explanation is not very convincing for the case of the weight guessing contest, as Scott Page remarks. This inadequacy does not seem to have struck Surowiecki, Sunstein, or any of those who have connected this and other similar anecdotes to the “miracle of aggregation.”

²⁸ 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).

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 more traditional explanation in terms of deliberation and the pursuit of rational consensus.²⁹ Compared to the Condorcet Jury Theorem, the miracle of aggregation presents the advantage of not making high demands on the average voter. In its elitist version especially, the cognitive threshold for the average voter can in particular be much lower than that of a coin flip.

Two main objections can be raised against that explanation.³⁰ 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). The other objection stems from Scott Page for whom an explanation that presupposes an infinity of independent signals is not plausible and who replaces it with an account of the wisdom of crowds in terms of aggregation of diverse predictive models, where a limited variety of perceived signals are interpreted differently by different individuals. Let me address these objections in turn.

The Problem of Voters’ “Rational Irrationality” and Systematic Cognitive Biases

²⁹ 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.

³⁰ A 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.

According to Caplan, the main problem with the miracle of aggregation scenario is that in order for the “miracle” to occur, one has to assume that there is a symmetrical distribution (random or else) of errors so that the mean of these errors is zero. Brian Caplan rejects this idea on the basis of evidence, essentially borrowed from psychology and behavioral economics (including his own research), showing that people are systematically biased in the same wrong direction. Caplan focuses on economic questions, where he thinks popular judgment is most obviously wrong but is willing to extend his claim to the entire political realm (and particularly to health policy and foreign policy). Caplan diagnoses four main economic misconceptions in the average citizen: 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 law is a good thing, free-trade is a positive sum game, 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.

When such individual biases are aggregated at the collective level and taken as the source of political decision, so Caplan argues, they are bound to produce drastically suboptimal policies. This leads Caplan to advocate for less democratic input and suggest a delegation of economic choices to those who know (economists) and, wherever possible, the market itself.

This criticism of voters’ systematic bias is very powerful and a priori convincing for the case of economic questions. Let me try however to raise a few (internal) objections. Ultimately I will turn to another explanation of collective wisdom than the miracle of aggregation but I want to suggest that it is less vulnerable to Caplan’s critique than it may seem.

First, one might object to the measure of voters’ “ignorance” and “incompetence” that Caplan subscribes to, namely the inability to answer questionnaires designed by political

scientists or economists. As argued in the first section, epistemic competence is poorly measured by this type of surveys so that the conclusions one can reach from observing low levels of information in voters are limited. The fact that Caplan correlates these low levels of information in voters with what he deems to be generally bad economic policies (such as economic protectionism) does not establish that the problem lies in voters' level of information. It might well be the case that voters are sufficiently informed about the consequences of economic protectionism and still choose it overall as what they think is the best policy, all things considered.

This is where Caplan can resort to another, independent argument, namely that of “systematic biases.” It is a fact that even informed voters can make mistakes due to cognitive biases and the use of (bad) heuristics. The gist of Caplan’s indictment lies in assuming that certain observed ideological preferences are on a par with such cognitive biases and heuristics. In the same way that people are known to suffer from base-rate neglect³¹ or to be subject to framing effects,³² they seem to be systematically anti-free-trade and pro-job security. All those biases make voters vote irrationally. One can here object that 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 (except economists) equally suffer from base-rate neglect, Americans are actually much less obsessed with job security than Europeans. The “systematic biases” denounced by Caplan are not as hard-wired as he sometimes seems to suggest.

Another objection that can be raised against Caplan’s anti-democratic conclusions bears on the assumptions on which his theory ultimately relies. On Caplan’s interpretation voters are rationally ideological (rationally “irrational”), since they chose to believe what

³¹ 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.

³² They give different answers to a same question that frames things differently.

makes them feel good rather than what is true. Caplan here reconciles classical rational choice theory and behavioral economics, giving an account of expressive or ideological voting in rational choice terms. This also implies, however, that people are just as interested as they are in the classical rational choice model. People simply maximize another form of self-interest than economic self-interest, namely the ideological pleasure derived from voting. The problem is that by voting ideologically, voters in fact fall prey of a form of irrationality—namely self-deception—since they could not get the benefits of their belief (the feel-good effect of altruism described by Caplan, or what Elster calls the “warm glow” effect) if they did not lie to themselves about their motivations. As a rich Hollywood producer voting for redistributive policies I can only feel good about myself if I think that my preference is genuinely altruistic. If I am aware that the only goal I am pursuing by voting left is not to help others, but to feel or appear generous, the pleasure I derive from my ideological vote disappears. So voters can only vote ideologically if they are both interested and irrational in some respect (self-deceptive)—two assumptions that do not square entirely with Caplan’s conclusions (that voters are rationally ideological).

If one gives up on the self-interest hypothesis, as this paper does by assuming that what rulers are pursuing is the common good, not their economic interest or their ideological pleasure, it must be the case that people not only vote altruistically, but do their best to vote smartly too. Altruism, after all, is only meaningful if it goes with some degree of rationality (wanting the means to one’s altruistic ends). I thus object to Caplan’s interpretation of citizens’ vote as merely ideological. If people are both disinterested and rational, they will not favor wrong beliefs over correct ones just because the former feel good. This is not to say that they will get out of their way to get more informed, but as I said earlier the impact of low information level on epistemic competence is unclear. Voters might be able to pass the right kind of judgment based on the little information they do have.

Finally, from a comparative point of view, one might object that the alternative envisaged by Caplan—rule of the economists—is not fool-proof either. The fact that professional economists do know better on a number of economic questions may seem to entitle them to rule unchallenged in that domain. Ideological thinking among experts, however, is as documented as voters' low level of information. Because of the lack of cognitive diversity that their group structurally presents, economist-kings are likely to deviate drastically from the right answer. Even granting the implausibility of the “miracle of aggregation,” absent a more general demonstration of the superiority of the small number over the larger number, all that Caplan achieves in demonstrating is that democracies can sometimes go wrong, not that another decision-procedure would do systematically better.

The other alternative to democracy proposed by Caplan is market law. Proposing the market as an alternative to democracy relies in my view on a confusion between decision on the one hand and information-aggregation or prediction on the other. The political alternatives to rule of the many are rule of one or the few, not the market. The market is just a tool available to all regimes. It might be better at aggregating information or making predictions than deliberation or voting, but unlike deliberation and voting (which also work as information-aggregating and predictive procedures), the market does not allow us to make any decisions. By definition the market operates where and when the political ends. But even in order to decide where the political ends, we need a decision-procedure—rule of one, few or many—not the market.

Let me now turn very briefly to Scott Page's critique. Scott Page dismisses both the Condorcet Jury Theorem and the Miracle of Aggregation for the same reason, namely on the grounds that they implausibly presuppose that voters receive an infinity of independent signals that they pick up in order to make a prediction. In reality 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 2006: 192)

3.3 The Crowd Beats the Average Law

Page proposes a different account of why large groups of people can make accurate predictions. 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 mistakes cancelling each other out. What matters 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. This account is not immune to the problem of systematic biases but given a large enough group and a liberal society encouraging dissent and diverse thinking, Caplan’s worst case scenario of a situation in which the average error is high and diversity low is not very plausible.

Page’s model also shows that the assumption of independence essential to the CJT and the miracle of explanation actually leads to understating the predictive ability of small groups and overstating the ability of large groups. His model, by getting rid of the assumption of independence, does the opposite. This slightly sobering from our democratic point of view, since this conclusion seems to favor oligarchy over democracy, but it is also enlightening as to the conditions under which large numbers of people can be smart. The existence and cultivation of the right kind of cognitive diversity in the group is one of them.

Let me now briefly present Page’s two main results: The Diversity Prediction Theorem and the Crowd Beats Average Law. I will not enter into the details of the demonstration behind either of them but take them as a starting point for my reflection on the epistemic properties of judgment aggregation through majority rule. The first theorem states

that a group's collective error equals the average individual error minus their collective predictive diversity (Page 2006: 208). In other words, when it comes to predicting outcomes (such as who is going to be a better president or whether a proposed document is the right kind of foundation for the European project, say), cognitive difference among voters matters just *as much as* individual ability. Increasing prediction diversity by a unit results in the same reduction in collective error as does increasing average ability by a unit.

Remember that in the case of deliberation applied to problem-solving, we saw that cognitive diversity could actually trump individual ability. Here the theorem demonstrates that there is a strict tie between these two components of collective intelligence. This is so because unlike what happens in deliberation the better information, idea, or argument does not crowd out the worse. Judgment aggregation aggregates everything, including the bad input, which occasionally makes the group less smart.

The second theorem states that 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 outpredicts its average member increases as the group becomes more diverse (Page 2006: 197). This "law" directly results from the Diversity Prediction Theorem. The implication for democracy is that we are better off making decisions that involve predictions through majority rule than we would be if the average member of the group made those decisions for us. Majority rule beats the rule of one (when the one is randomly chosen). Notice, however, that because diversity does not trump ability when it comes to predictions, we might be just as well as a group if a smaller group of smarter people made the decisions. Majority rule among the many does not beat majority rule among the few.³³

³³ Page, however, also establishes another theorem, the "Crowd's Possibly Free Lunch Theorem", according to which, roughly speaking, it can happen that a group of random people with diverse simple predictive models based on different interpretations of the world can predict a complicated function, which beats in accuracy that designed by expert statisticians. An example might help. Let us say the goal is to figure out whether an

What is the import of Page's findings for the idea of democratic reason? Majority rule, and in fact any democratic mechanism that aggregates individual judgments into collective judgments do have epistemic properties. Since the group's prediction beats that of the average citizen in the group, we have an argument why the rule of the many beats 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 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.

4. Conclusion

Combining the epistemic properties of deliberation and majority rule, I conclude that democracy beats any version of the rule of the few, including when we make unrealistic assumptions about the intelligence of the few. 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 though 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, apparently at least, for democracy in the second case. Holding both the virtue and information factors constant, the rule of the many beats the rule of the few. Let us now consider what happens when we (theoretically) reintroduce those two variables.

incumbent to the presidency will be re-elected. Obviously there are multiple factors to take into account to make that prediction. Most of us, however, will just rely on just one or two variables. I will use, say, the unemployment rate. You will use the age of the incumbent and the level of oil prices. If we aggregate many people's predictive and sufficiently different models, they can occasionally add up to an extremely complicated function that could have been designed by a group of experts (except that the coefficients are not likely to be very accurate). If the group is lucky, this function yields more accurate predictions than those devised by experts themselves. In Page's words, "simple, diverse predictive models can form sophisticated crowd-level predictions" and "these crowds can perform better than experts, provided their increased coverage more than makes up for the crudeness of their estimates" (Page 2006: 234). Page admits that such a "free lunch" is unlikely but possible.

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 an oligarchy not to abuse an unchallenged power to do what can best serve them, even if they have to make some concessions to the masses to keep them quiet. By contrast, the rule of the many, which is structurally designed to rule for the greatest number, economizes on virtue as it economizes on intelligence. To the extent that collective epistemic competence is also a function of the decision-makers' virtue, democracy a fortiori beats the rule of the few when we reintroduce the virtue component.

If we now reintroduce the information component, 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. I assume that this applies to the decisions of representatives too, given mechanisms of accountability that more or less tie the judgments of citizens to the decisions of those they elect (following Caplan 2007 on both points here). Even admitting that this assumption is correct, it remains an open question whether low levels of information directly translate into low epistemic competence. We know too little still about the complex ways in which voters make up for lack of information to be able to conclude that democratic citizens are definitely unable to pass enlightened judgments based on little information. The democratic mechanisms of deliberation and voting might be precisely why they need not become more informed individually, if those mechanisms are able, as I argue, to turn their relatively weak input into a much better output.

On the other hand, we saw that the great property of democratic deliberation is to bring out the relevant information more efficiently than deliberation among the few. 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. Notice that here I am talking about information that has direct relevance for the question at hand, since it is brought up in context, not the type of information measured by empirical surveys. Whether the amount of information made collectively available through democratic deliberation more than compensates the disincentives to get informed that citizens have in voting 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 an 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 us just suggest that whatever might be, or might have been, the initial reasons to prefer democracy over dictatorship or oligarchy, collective wisdom, what I call democratic reason, might help explain why we keep it.³⁴

Let me add a final word on the conditions for democratic reason to emerge. I have insisted, following Page, 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, which might be called, broadly speaking, “liberal.” This society must be 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

³⁴ Josiah Ober’s manuscript on “Democracy and Knowledge” (to appear in 2009) confirms this idea, proposing (from what I understand) on the basis of empirical evidence that the superiority of Athens over rival city-states came from the epistemic properties of its democratic institutions. I came across that manuscript only recently and still have to figure out more precisely how his thesis bears on the epistemic case for democracy.

democratic reason is thus conditional on the existence of a social and cultural context that nurtures and protects, among other differences, cognitive differences.

To the extent that the epistemic argument for democracy has value, therefore, it establishes that democracy and liberalism go together. In other words, democracy is more likely to be smart only if it is, also, liberal. 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.

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.

Cohen, Joshua. 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.

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 and (eds). 2008 forthcoming. *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.

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.
- Goodwin 1992. REF
- Gottfredson, Linda S. 1997. “Why g matters: the complexity of everyday life.” *Intelligence* 24 (1):
- 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.
- Lave, Jean. 1988. *Cognition in Practice: Mind, Mathematics, and Culture in Everyday Life*. New York, NY: Cambridge University Press.

- 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.
- List, Christian and Robert E. Goodin. 2001 "Epistemic Democracy: Generalizing the Condorcet Jury Theorem." *Journal of Political Philosophy* 9 (3), 227-306. How elitism undermines the study of voter competence
- Lupia, Arthur. 2006. "How Elitism undermines the study of voter competence." *Critical Review* 18 (1-3) :217-232.
- Mackie, Gerry. 2002. "Does Democratic Deliberation Change Minds?" prepared for the Workshop, Democratic Theory: The Canberra Papers, 27 March 2002.
- Manin, Bernard. 1996. *Les Principes du Gouvernement Représentatif*. REF
- 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. "Cognivite Artifacts." In Carroll JM (ed) *Designing Interaction: Psychology at the Human-Computer Interface*: 17-38. New York, NY: Cambridge University Press.
- Ober, Josiah. Forthcoming 2009. *Democracy and Knowledge*.
- 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.
- 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 an 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*. REF.

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.

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. 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, JV. 1998. *Mind as Action*. New York, NY: Oxford University Press