It's the Aggregation, Stupid!

Josh Chafetz


"Men, it has been well said, think in herds; it will be seen that they go mad in herds, while they only recover their senses slowly, and one by one."¹ So wrote Charles Mackay in the Preface to his Extraordinary Popular Delusions and the Madness of Crowds, a work that aimed to catalogue mass delusions from the Crusades to witch hunts, the South Sea bubble to the tulip craze, alchemy to fortune telling. Mackay's study of the irrational behavior of crowds was unusually detailed, but it hardly stands alone. Over two thousand years earlier, Socrates had complained to Crito, "Would that the majority could inflict the greatest evils, for they would then be capable of the greatest good, and that would be fine, but now they cannot do either. They cannot make a man either wise or foolish, but they inflict things haphazardly."² Nineteenth-century social theorist Gustave Le Bon was even more dismissive:

This very fact that crowds possess in common ordinary qualities explains why they can never accomplish acts demanding a high degree of intelligence. The decisions affecting matters of general interest come to by an assembly of men of distinction, but specialists in different walks of life, are not sensibly superior to the decisions that would be adopted by a gathering of imbeciles. The truth is, they can only bring to bear in common on the work in hand those mediocre qualities which are the birthright of every average individual. In crowds it is stupidity and not mother-wit that is accumulated.³

And Nietzsche, in characteristically pithy style, wrote that, "Madness is rare in individuals—but in groups, parties, nations, and ages it is the rule."⁴ Indeed,

² Staff Writer, The New Yorker.
⁴ Plato, Crito, in FIVE DIALOGUES 45, 47 (G.M.A. Grube trans., Hackett Publ’g Co. 2d ed. 2002).
so widely accepted is the proposition that crowd psychology is pathological that it seems hardly to require supporting argumentation. Mackay’s “madness of crowds” thesis has become almost axiomatic.\(^5\)

In economic theory, however, there has long been an understanding, expressed most eloquently in Hayek’s discussion of spontaneous order, that crowds can often make better decisions than any individual.\(^6\) The simplest example is the price system. Every consumer has different preferences as to what kinds of food he likes. Consumers themselves may not even be conscious of their rank-ordering of foods—perhaps it only manifests itself unconsciously in their shopping patterns. Certainly, a central planner could not gather all of this information—even if the consumers were conscious of their preferences, the planner would have to be continuously polling consumers; and if they were not fully conscious of their preferences, then there would be no effective means of centrally gathering this information. However, in a free market, it is unusual to go to the store and discover that the type of food one wants is out of stock. This is because the price structure serves as a signaling mechanism. If many consumers prefer beef to chicken, the price of beef will rise, and the price of chicken will fall. This signals producers to supply more beef and less chicken. As supply adjusts in response to demand, the prices and quantities supplied of beef and chicken stabilize at the level at which consumers get what they want. The suppliers of beef and chicken have not taken polls of personal preferences, nor have they arranged with one another to inversely adjust their production. Instead, they have simply looked at market-produced signals—prices—and set their levels of output so as to maximize their own profits. The aggregated effect of these individual actions produces a better result than any single mind could have. Borrowing language from Adam Ferguson,\(^7\) Hayek refers to “orderly structures which are the product of the action of many men but are not the result of human design.”\(^8\) That is, an orderly structure far too complex for a human mind to have designed—an economic system—has been created by the behavior of a crowd.\(^9\) In Le Bon’s terms, price serves as a mechanism to accumulate, not the crowd’s stupidity, but its mother-wit.

Part of James Surowiecki’s goal in *The Wisdom of Crowds* is to show that

---

\(^5\) Crowd pathology is, of course, also a mainstay of literature and pop culture. See, e.g., HENRIK IBSEN, *An Enemy of the People*, in *FOUR GREAT PLAYS BY IBSEN* 129 (R. Farquharson Sharp trans., Bantam Books 1959) (1882); ARTHUR MILLER, *The Crucible* (1952); PHILIP ROTH, *The Plot Against America* (2004); WILLIAM SHAKESPEARE, *The Tragedy of Julius Caesar* act 1, sc. 1; MONTY PYTHON’S *LIFE OF BRIAN* (Handmade Films 1979).


\(^7\) See ADAM FERGUSON, *AN ESSAY ON THE HISTORY OF CIVIL SOCIETY* 119 (Fania Oz-Salzberger ed., Cambridge Univ. Press 1995) (1767) (observing that “nations stumble upon establishments, which are indeed the result of human action, but not the execution of any human design”).

\(^8\) 1 HAYEK, * supra* note 6, at 37.

\(^9\) See 3 F. A. HAYEK, *LAW, LEGISLATION AND LIBERTY: THE POLITICAL ORDER OF A FREE PEOPLE* 164 (1979) (“We have never designed our economic system. We were not intelligent enough for that.”).
this Hayekian logic extends far beyond the economic field. When asked to
guess the weight of an animal at a livestock show, the right answer on Who
Wants to Be a Millionaire?, or the results of an election, the average guess
of a crowd is often at least as good, if not better, than the best guess of any
member of the crowd. But Surowiecki is not blind to the fact that crowds can
also make bad decisions—worse, sometimes, than even their dumbest
component parts. In that vein, he discusses stock market bubbles, traffic
jams, and the NASA deliberations which concluded that nothing could or
should be done to repair the Columbia's wing.

Of course, the truly interesting questions concern how we distinguish the
good crowds from the bad ones, and what can be done to turn the latter into the
former. It is on this point that Surowiecki makes his sharpest contribution: He
identifies four criteria of wise crowds. First, they have a diversity of opinion:
"[E]ach person should have some private information, even if it's just an
eccentric interpretation of the known facts." Second, the members should be
independent of one another—their opinions should not be determined by the
opinions of those around them. Third, the crowd's decisionmaking should be
decentralized—each individual should be able to make use of the local
knowledge available to her. Finally, there must exist some good mechanism for
aggregating the judgments of the members of the crowd into a collective
decision.

Although Surowiecki never puts it in these terms, the basic insight seems to
be this: Every individual decision is part signal and part noise. As long as the
noise is randomly distributed around the signal, then the signals of the group
members will add together, while the noise will cancel out. Viewing
Surowiecki's four criteria in these terms, diversity is meant to ensure that the
signal contributed by each group member is not identical to the signal
contributed by other group members. Independence and decentralization are
meant to protect against a systematic bias in the distribution of noise by
ensuring that group members do not adopt the noise of others as a result of
pressure, the desire to conform, or a lack of epistemic confidence. And an
appropriate aggregating mechanism is necessary to make sure that the noise
really does cancel out.

11. Id. at 3-4.
12. Id. at 17-19.
13. Id. at 241-58.
14. Id. at 145-57.
15. Id. at 173-91.
16. Id. at 10.
17. Id.
18. To illustrate: If people are asked to guess the number of jellybeans in a jar, each guess will be
based on some information (an observation of the jar) and some uncertainty (an inability to count
precisely). So long as there is nothing causing a systematic bias (i.e., causing people to tend to
Using his four criteria of wise crowds, Surowiecki can also explain how crowds can go very wrong. For example, he views stock market bubbles as instances of excessive dependence among investors, a violation of his second criterion of wise crowds. A bubble results when investors stop believing their own independent judgments of the value of a stock and instead assume that, because everyone else seems to think it is worth more, it must be worth more. Likewise, traffic jams can be seen as failures of aggregation (the fourth criterion)—drivers lack a means, akin to the role of price in the market, of aggregating their preferences in such a way that only those to whom it is worth the most drive during rush hour. And NASA’s tragic failure to do anything about the Columbia problem resulted from a lack of independence and too centralized a decisionmaking process (the pressures for groupthink within the Mission Management Team were strong), violating Surowiecki’s second and third criteria. Surowiecki also suggests that NASA is insufficiently diverse (the first criterion)—unlike the NASA of the 1960s, “NASA employees today are far more likely to have come to the agency directly out of graduate school, which means they are also far less likely to have divergent opinions.”

Surowiecki’s work is obviously relevant to a number of legal and policy fields, but it has perhaps the most interesting implications for two: juries and theories of deliberative democracy. Both the jury system and theories of democracy presuppose that crowds can and generally will make good decisions. Deliberative democracy—a vision of democratic governance that gives a central place to the public deliberation of the citizenry—is the democratic theory that focuses most closely on group decisionmaking. Since Surowiecki’s book speaks precisely to the question of how well groups can make decisions, it is worth pondering what it can teach us about juries and overestimate or underestimate), the uncertainty (“noise”) will fall roughly equally on both sides of the correct answer. The simple aggregating mechanism of taking the arithmetic mean of the guesses will thus cause the noise to cancel out, and the mean will tend to be very close to the correct answer. Indeed, the mean is often closer than even the best individual guess, and—if the experiment is repeated several times—the mean will almost certainly be better than even the best individual guess. Id. at 5. What is true of a simple factual issue like counting jellybeans is also true of a more complex coordination problem, such as the decision whether to produce beef or chicken discussed supra notes 6-9 and accompanying text. No producer has a full picture of the market—each has some information (signal) and some guesswork or forecasts (noise). By rewarding those whose guesses are in line with consumers’ wants and punishing those whose guesses are not, price serves as an aggregating mechanism to produce a better outcome than would be produced by any individual decisionmaker.

19. Id. at 248.

20. But see Emma Smith, Coming Soon to a City Near You: The Congestion Charge, SUNDAY TIMES (London), Feb. 6, 2005, at 6 (noting the success of London’s “congestion charge,” in which drivers pay a fee to drive during the day in parts of central London, and noting plans to introduce a similar charge in other cities). Surowiecki cites the London congestion charge approvingly. SUROWIECKI, supra note 10, at 145-50.

21. SUROWIECKI, supra note 10, at 183.

22. For a fascinating case study that cannot be discussed in detail in this Comment, see id. at 79-83 (discussing FutureMAP, the Defense Advanced Research Projects Agency’s much-maligned, and ultimately abandoned, attempt to set up a futures market in forecasting events relevant to national security).
The Wisdom of Crowds
theories of deliberative democracy.

A jury, like any small deliberative group, presents the danger that individuals' opinions will be too heavily influenced by the opinions of others, and that the decisionmaking process will become too centralized. Surowiecki warns of both group polarization and the exercise of undue influence by especially powerful or talkative members of the group.23 These are not merely speculative worries. As Cass Sunstein has shown, jury-like deliberation by a group of ideologically similar people tends to amplify the ideological predisposition of the group's members. In other words, if the members of a group were predisposed toward a severe punishment going in, the result of the deliberation will often be more severe than the median punishment preference before deliberation.24

But if Surowiecki's criteria for successful group decisionmaking demonstrate some potential pitfalls of juries, they also point the way towards solutions. As he notes, "groups can be, as it were, depolarized."25 Studies have suggested that the presence of even a single dissenter may significantly improve the quality of the result, both by acting as a brake on the incestuous amplification that leads to increasing polarization and by emboldening others who disagree with what seems to be the dominant view to speak up.26 It is possible that fairly minor reforms could help dissent emerge in the jury room. For example, suppose that jurors were each asked to write down a few key thoughts about the case and to begin joint deliberations by going around the table and reading out what they had written. Since they will have written their thoughts before they know what the dominant ideas amongst the other jurors are, dissent will be more likely to emerge. And, as Sunstein has observed, once some dissent emerges, others will be emboldened to speak their minds more freely as well.27 That is, the introduction of divergent viewpoints from the beginning will increase the overall quality of the deliberation by increasing diversity and decreasing dependence and centralization. Along the same lines, it might be worth considering ways that judges could actively foster greater viewpoint diversity in the selection of jury panels.28

23. Id. at 184-88.
25. Surowiecki, supra note 10, at 188.
27. See id.
28. It is hardly a new idea that judges can seek to foster racial diversity on jury panels. See Nancy S. Marder, Juries, Justice & Multiculturalism, 75 S. Cal. L. Rev. 659 (2002); Joshua Wilkenfeld, Note, Newly Compelling: Reexamining Judicial Construction of Juries in the Aftermath of Grutter v. Bollinger, 104 Colum. L. Rev. 2291 (2004). Whether such action by judges violates the Equal Protection Clause is well beyond the scope of this Comment. Although at least one federal appellate court has held that any racial (or religion-based) "jurymandering" is unconstitutional, see United States
There remains the question of aggregation in juries. In criminal trials, we tend to think that juries should be unanimous, and in civil trials we expect at least a supermajority. These requirements will naturally suggest a consensus-driven process in the jury room. But, as Surowiecki notes, "the search for consensus encourages tepid, lowest-common-denominator solutions which offend no one rather than exciting everyone. . . . [C]onsensus-driven groups—especially when the members are familiar with each other—tend to trade in the familiar and squelch provocative debate." jurors, of course, are not familiar with each other beforehand. Moreover, as noted above, we can seek ways to encourage the expression of dissent in the jury room and, to that extent, mitigate the dangers of a consensus-based process. For binary jury decisions (guilty/not guilty, liable/not liable), this may be the best we can do. For non-binary decisions (damage amounts, sentence lengths, perhaps even degrees of murder), other aggregation techniques—perhaps an arithmetic mean—may be preferable.

Insofar as Surowiecki’s book provides lessons applicable to the jury system, so too it speaks to theories of deliberative democracy, which can be seen as an attempt to govern by juries writ large. David Miller has identified the “deliberative ideal” as involving “an open and uncoerced discussion of the issue at stake with the aim of arriving at an agreed judgement. The process of reaching a decision will also be a process whereby initial preferences are transformed to take account of the views of others.” Joshua Cohen has given a more specific vision, laying out four principles of “ideal deliberation”: (1) deliberation must be free, in that participants regard themselves as bound only by the preconditions and results of their deliberation and suppose that they are

v. Nelson, 277 F.3d 164, 207-10 (2d Cir. 2002), commentators have questioned whether this holding remains good law. See Wilkenfeld, supra (arguing that Nelson is no longer good law in light of Grutter v. Bollinger, 539 U.S. 306 (2003)). Although the constitutionality of a legitimate attempt to foster a diversity of viewpoints on jury panels is not entirely clear, it would seem to present less of an equal protection problem than attempts to foster racial diversity.

29. But see the Supreme Court’s decisions in the companion cases of Johnson v. Louisiana, 406 U.S. 356 (1972), and Apodaca v. Oregon, 406 U.S. 404 (1972). Justice Powell, who provided the controlling opinion in these cases, agreed with the Johnson dissenters that the Sixth Amendment requires jury unanimity in federal criminal trials, but argued that this requirement has not been incorporated into the Fourteenth Amendment and hence is not applicable in state criminal proceedings. Johnson, 406 U.S. at 366 (Powell, J., concurring).

30. SUROWIECKI, supra note 10, at 203.

31. Unfortunately, this practice is currently disallowed in most jurisdictions. See 75B AM. JUR. 2D Trial § 1782, at 544 (1992) (“A quotient verdict arrived at after an agreement by jurors to arrive at the amount of the verdict by aggregating the vote of all jurors and dividing the result by 12 is generally void.”); id. at 546 (“The rule as to quotient verdicts is applicable to both criminal and civil cases alike. Thus, where the jury in a criminal case fixes the term of imprisonment by setting down the periods which each juror favors and dividing the aggregate by 12, agreeing in advance to abide by the result, the verdict will be held invalid.”); see also id. at 544 n.96 (citing cases disapproving quotient damage calculations in civil actions); id. at 546 n.6 (citing cases disapproving quotient sentence length calculations in criminal actions).

32. David Miller, Deliberative Democracy and Social Choice, POL. STUD., Special Issue 1992, at 54, 55 (citation omitted).
The Wisdom of Crowds

able to act from their results; (2) deliberation must be reasoned, in that participants are expected to advance justifications for their proposals; (3) deliberation must be characterized by equality among the participants; and (4) deliberation must seek to arrive at consensus, with a vote being taken only when consensus proves impossible. Cohen also insists that the deliberating body be pluralistic, meaning that its members have “diverse preferences, convictions and ideals concerning the conduct of their own lives.”

Applying Surowiecki’s criteria to this conception of democracy is illuminating. As we have already seen, the diversity requirement is important. A non-diverse deliberative body—like a non-diverse jury—might well reach a result more extreme than the median preference held by those deliberators prior to deliberations. Proposals that involve neighborhood-based deliberative groupings run a serious risk of incurring this sort of incestuous amplification in relatively homogenous neighborhoods. Deliberation also obviously poses threats to the independence and decentralization of the decisionmaking processes of group members. There is a risk that individuals will feel pressure to go along with what seems to be the dominant opinion, a risk that is heightened if the deliberating group is not very diverse. The use of consensus as an aggregation mechanism pushes further in this direction, as we saw with juries. It is not clear that the use of a “moderator” would solve these problems. A moderator could only ensure that those who want to speak have access to the floor; she would have no way of convincing the reticent to speak or of preventing polarization. It is also worth noting that the presence of a moderator cuts against Cohen’s equality criterion for ideal deliberation. It seems likely that a moderator would have disproportionate influence over the deliberations and that unconscious biases would affect the way the moderator treats the deliberators. A proposal to replace the moderator with a “citizen-foreman” is unavailing. As Surowiecki notes, two-thirds of all jury foremen

34. Id. at 21.
35. See, e.g., BRUCE ACKERMAN & JAMES S. FISHKIN, DELIBERATION DAY 24 (2004) (“After arriving at neighborhood schools and community centers between 8 and 9 A.M., deliberators will be randomly assigned to groups of fifteen for the first event...” (emphasis added)).
36. Ackerman and Fishkin point to the success of Fishkin’s Deliberative Polls as evidence that deliberative groups will not polarize. Id. at 3-5, 62-65. But Fishkin’s Polls have intentionally gathered representative—that is, diverse—samples of the general population. See id. at 4; Sunstein, supra note 24, at 117 (“Fishkin’s groups were highly diverse and enclave deliberation was impossible.”). The lack of polarization in Fishkin’s samples can, as we have seen, be attributed to their diversity. Ackerman and Fishkin could attempt to remedy this problem by busing people around on Deliberation Day to interject diverse viewpoints into the deliberations of homogenous neighborhoods, but (in addition to the expense and inconvenience for all involved) this would make deliberation on local issues and local elections impossible.
37. See ACKERMAN & FISHKIN, supra note 35, at 48, 69.
38. Id. at 66.
are men, and there is no reason to think that the foreman selection process is not biased in other ways as well. Ackerman and Fishkin propose to remedy this by selecting foremen by lot, with an opt-out provision. It seems likely, however, that those most reticent to speak up during deliberations would also be most likely to opt out of the foreman selection process; that is, the foreman selection process may conduce to polarization rather than inhibit it. Moreover, a citizen-foreman presents one problem that a moderator does not. Presumably, a moderator stands outside the debate—her function is solely to moderate, not to participate. But a foreman deliberates along with the group. Now, a **member** of the group—one who presumably has the right to speak on substantive matters like everyone else—has been elevated in status, thus doing further damage to the equality principle in deliberation.

Does this mean that we should give up on deliberation and return to a Schumpeterian minimalist conception of democracy? That conclusion would be premature. Organized deliberation may well be desirable in many cases; Surowiecki’s criteria for effective group decisionmaking may simply help us better organize deliberative events. First, we return to the importance of the diversity of the deliberating body. Because non-diverse groups tend to polarize in unhealthy ways, it may be preferable to forego organized deliberation entirely in favor of simply aggregating the initial judgments of group members in cases in which a diverse group is impossible to assemble. Where it is possible to assemble a diverse deliberative group, strong efforts should be made to do so. Second, the deliberations must be structured so as to foster independent decisionmaking. Insofar as a moderator or foreman is necessary, she must be given as little discretion as possible. The moderator or foreman, however, is not the only potential threat to independence. It will be more difficult to guard against the possibility that an overbearing speaker or a powerful local community member will overawe his fellow deliberators. It is not clear that any institutional design can prevent this; the success of deliberations will depend to a large degree on the development of social norms of deliberative equality and respect for dissent. Finally, and perhaps most importantly, we return to the aggregation mechanism. A strong norm in favor of consensus acts both as a deterrent to dissent (who wants to be the one person...
preventing the entire community from reaching a decision?) and as a constraint on the types of ideas even worth presenting (why go to the trouble of explaining your unconventional stance when it has no chance of gaining a consensus?). Deliberation without the need to come to a consensus may be an excellent way of putting the pros and cons of many ideas on the table; a consensus requirement, however, may tend to constrain the field of ideas in play and stifle dissenting voices. It is thus one of the great virtues of Ackerman and Fishkin’s proposal that final decisions are not made at the end of the day—the citizen goes home, ponders what she has learned, and goes to the polls, as an individual, two weeks later. She is thus free to aggregate the information she has received, from organized deliberation and elsewhere, to weigh it according to her own unique values, preferences, and interpretations, to take advantage of her local knowledge, and to vote independently. There will still be some noise—some voter ignorance—in every vote, but as long as the sample is large enough and there is no systematic bias, the ignorance will scatter more or less randomly around the voting options. In the end, it may be the ability to aggregate individual votes that saves the concept of group deliberation.

*The Wisdom of Crowds* is an impressive little book. Not only does it take the theory that crowds can be excellent decisionmakers out of the realm of abstract economic theory and bring it to a popular audience, but it also compiles a diverse and fascinating set of examples of good (and bad) group decisionmaking. Most interestingly, it advances an analytically powerful set of criteria for good group decisionmaking. As we’ve seen, the book’s thesis can help us to evaluate and refine the institutional structures for juries and deliberative democracy. But Surowiecki’s book has implications still further afield, and it is to be hoped that its lessons will be contemplated in a wide array of disciplines.

---

43. *Id.* at 3.

44. It is worth noting that there may be other justifications for deliberation, unrelated to its efficacy in producing the best answer to any particular question. Political and legal losers may be more likely to accept the legitimacy of a system in which decisions are made by deliberation and consensus rather than on a straight up-or-down vote. And deliberation itself may conduce to social cohesion in a way that we find desirable. See, e.g., ALEXIS DE TOCQUEVILLE, *DEMOCRACY IN AMERICA* 270-76 (J.P. Mayer ed. & George Lawrence trans., Doubleday 1969) (1850) (discussing various civic-republican functions served by juries). Additionally, some may think that deliberation on normative issues serves an inherently legitimizing social function. See JÜRGEN HABERMAS, *BETWEEN FACTS AND NORMS: CONTRIBUTIONS TO A DISCOURSE THEORY OF LAW AND DEMOCRACY* 287-328 (William Rehg trans., MIT Press 1996) (1992). The force of these justifications is beyond the scope of this Comment.