GROUP JUDGMENTS: STATISTICAL MEANS, DELIBERATION, AND INFORMATION MARKETS

CASS R. SUNSTEIN*

How can groups elicit and aggregate the information held by their individual members? There are three possibilities. Groups might use the statistical mean of individual judgments; they might encourage deliberation; or they might use information markets. In both private and public institutions, deliberation is the standard way of proceeding; but for two reasons, deliberating groups often fail to make good decisions. First, the statements and acts of some group members convey relevant information, and that information often leads other people not to disclose what they know. Second, social pressures, imposed by some group members, often lead other group members to silence themselves because of fear of disapproval and associated harms. As a result, deliberation often produces a series of unfortunate results: the amplification of errors, hidden profiles, cascade effects, and group polarization. A variety of steps can be taken to ensure that deliberating groups obtain the information held by their members; restructuring private incentives, in a way that increases disclosure, is the place to start. Information markets have substantial advantages over group deliberation; such markets count among the most intriguing institutional innovations of the last quarter-century and should be used far more frequently than they now are. One advantage of information markets is that they tend to correct, rather than to amplify, the effects of individual errors. Another advantage is that they create powerful incentives to disclose, rather than to conceal, privately held information. Information markets thus provide the basis for a Hayekian critique of many current celebrations of political deliberation. They also provide a valuable heuristic for understanding how to make deliberation work better. These points bear on the discussion of normative issues, in which deliberation might also fail to improve group thinking, and in which identifiable reforms could produce better outcomes. Applications include the behavior of juries, multimember judicial panels, administrative agencies, and congressional committees; analogies, also involving information aggregation, include open source software, Internet "wikis," and weblogs.

---

* Karl N. Llewellyn Distinguished Service Professor of Jurisprudence, Law School and Department of Political Science, the University of Chicago. For valuable comments, I am grateful to Michael Abramowicz, Bruce Ackerman, Ian Ayres, Jon Elster, Bernard Harcourt, Saul Levmore, Douglas Lichtman, Richard Posner, and Adam Samaha. Versions of this Article were presented as the Harold Leventhal Memorial Lecture at Columbia Law School and as the keynote lecture on Whither Democracy? at Brandeis University, and I am very grateful for the comments and suggestions offered on those occasions. I am also grateful to participants in workshops at the Brookings-American Enterprise Institute Joint Center on Regulatory Policy, Harvard Law School, the John F. Kennedy School of Government at Harvard University, the University of Chicago Law School, and the Yale Law School. Many thanks also to Robert Park for extraordinary research assistance. Thanks above all to Reid Hastie, for a great deal of patient help and tutoring.

Imaged with Permission from N.Y.U. Law Review

"Increased accuracy is a common justification for using groups, rather than individuals, to make judgments. However, the empirical literature shows that groups excel as judges only under limited conditions. . . . [G]roups performing tasks that involve solutions that are not easily demonstrable tend to perform at the level of their average members."[1]

"The presumption that Iraq had active WMD programs was so strong that formalized [Intelligence Community] mechanisms established to challenge assumptions and 'group think,' such as 'red teams,' 'devil's advocacy,' and other types of alternative or competitive analysis, were not utilized."[2]

"Sometimes important forecasts are made in traditional group meetings. This . . . should be avoided because it does not use information efficiently. A structured approach for combining independent forecasts is invariably more accurate."[3]

INTRODUCTION

How can groups obtain the information that their members have? There are three principal answers. First, groups might use the statistical mean (or median) of the independent judgments of their members. Second, groups might attempt to improve on those independent judgments by using deliberation and asking for the reasoned exchange of facts, ideas, and opinions. Third, groups might use information markets, through which group members, or those outside of the group, "bet" on their judgments about future events. Of course each of these methods can take diverse forms; one of my principal goals here is to explore which forms are most likely to produce good outcomes. The choice has implications for many institutions involved in law and politics, including the White House, juries, administrative agencies, congressional committees, federal courts of appeals, and even the Supreme Court itself.

Both private and public institutions tend not to rely on statistical means, and they rarely enlist information markets. Such institutions typically prefer to make decisions through deliberation. Generalizing from this fact, many people have paid a great deal of attention to

---

[1] Daniel Gigone & Reid Hastie, *Proper Analysis of the Accuracy of Group Judgments*, 121 PSYCHOL. BULL. 149, 149 (1997) [hereinafter Gigone & Hastie, *Proper Analysis*].

[2] SENATE SELECT COMM. ON INTELLIGENCE, REPORT OF THE 108TH CONGRESS, U.S. INTELLIGENCE COMMUNITY'S PREWAR INTELLIGENCE ASSESSMENTS ON IRAQ: CONCLUSIONS 7 (full version, S. REP. NO. 108-301 (2004)), *available at* http://intelligence.senate.gov.

[3] J. Scott Armstrong, *Combining Forecasts*, *in* PRINCIPLES OF FORECASTING 417, 433 (J. Scott Armstrong ed., 2001).

Imaged with Permission from N.Y.U. Law Review

deliberative accounts of democracy itself. The theoretical foundations of deliberative democracy have been elaborated in some detail,[4] and increasing attention is being devoted to methods for making democratic processes more deliberative. James Fishkin, for example, has pioneered the idea of the "deliberative poll," by which people are asked to deliberate together on public issues and to state their judgments only after the deliberative process is complete.[5] Fishkin and Bruce Ackerman have gone so far as to suggest a new national holiday, Deliberation Day, on which people are asked to congregate in groups in order to discuss and debate important issues of public policy.[6] Perhaps the proposal is unrealistic; perhaps citizens as a whole should not be expected to deliberate much in a liberal society.[7] But even if this is true, leaders in the public and private sphere might be urged to deliberate more than they now do, and many accounts of deliberative democracy emphasize the importance of deliberation by representatives.[8]

Why, exactly, is deliberation important or even desirable? A central answer must be that deliberation will result in wiser judgments and better outcomes.[9] But does deliberation actually have this effect? The answer is by no means clear.[10] In fact, group members may

---

[4] *See generally* DELIBERATIVE DEMOCRACY (Jon Elster ed., 1998) (collecting diverse treatments of deliberative democracy); AMY GUTMANN & DENNIS THOMPSON, DEMOCRACY AND DISAGREEMENT (1996) (defending deliberative democracy and discussing its preconditions); JÜRGEN HABERMAS, BETWEEN FACTS AND NORMS (1998) (elaborating deliberative conception of democracy). On the role of deliberative democracy in the American framing, see JOSEPH M. BESSETTE, THE MILD VOICE OF REASON 6–39 (1994), contending that framers contemplated deliberative democracy.

[5] *See* JAMES S. FISHKIN, THE VOICE OF THE PEOPLE 162–76 (1995).

[6] *See* BRUCE ACKERMAN & JAMES S. FISHKIN, DELIBERATION DAY 3–16 (2004) (proposing national day for political deliberation).

[7] *See* RICHARD A. POSNER, LAW, PRAGMATISM, AND DEMOCRACY 134–43 (2003) (challenging deliberative democracy as unrealistic).

[8] *See* BESSETTE, *supra* note 4, at 1–2 (discussing founders' emphasis on deliberating through representatives).

[9] There are other possibilities, of course. Perhaps deliberation has educative effects or contributes to individual self-development. Perhaps it legitimates decisions or increases the likelihood that people will acquiesce in them. *Cf.* TOM R. TYLER, WHY PEOPLE OBEY THE LAW 115–24 (1990) (discussing effects of fair procedures on legitimacy of and compliance with legal decisions). I am mostly putting these arguments to one side and focusing on the possibility that deliberation will improve outcomes. As noted below, however, deliberation tends to increase confidence and to decrease variance, even when it does not increase accuracy; it follows that deliberation might be justified because of its legitimating effects even when it fails to produce better outcomes. *See infra* text accompanying notes 77–84.

[10] *See* Robert J. MacCoun, *Comparing Micro and Macro Rationality*, *in* JUDGMENTS, DECISIONS, AND PUBLIC POLICY 116, 121–26 (Rajeev Gowda & Jeffrey C. Fox eds., 2002) (discussing factors that can cause groups to judge less accurately than individuals); Gigone & Hastie, *Proper Analysis*, *supra* note 1, at 161–62 (discussing how group judgments are

Imaged with Permission from N.Y.U. Law Review

impose pressures on one another, leading to a consensus on falsehood rather than truth. The idea of "groupthink," coined and elaborated by Irving Janis, suggests the possibility that groups will tend toward uniformity and censorship, thus failing to combine information and enlarge the range of arguments.[11] Without structural protections, both private and public groups may well err, not in spite of deliberation but because of it. By contrast, the use of statistical means or of information markets will often lead to more accurate decisions.

As an example of a failure of deliberation, consider the account in the 2004 report of the Senate Select Committee on Intelligence. That report explicitly accused the CIA of groupthink, through which the agency's predisposition to find a serious threat from Iraq caused it to fail to explore alternative possibilities or to obtain and use the information that it actually held.[12] In the Committee's view, the CIA "demonstrated several aspects of group think: examining few alternatives, selective gathering of information, pressure to conform within the group or withhold criticism, and collective rationalization."[13] Thus the agency showed a "tendency to reject information that contradicted the presumption" that Iraq had weapons of mass destruction.[14] Because of that presumption, the agency failed to use its own formalized methods "to challenge assumptions and 'group think,' such as 'red teams,' 'devil's advocacy,' and other types of alternative or competitive analysis."[15] Above all, the Committee's conclusions emphasize the CIA's failure to elicit and aggregate information that was actually in the possession of its employees.

This claim is a remarkable and even uncanny echo of one that followed the 2003 investigation of failures at NASA surrounding the explosion of the space shuttle Columbia. The investigation stressed the agency's similar failure to elicit competing views, including those based on information held by agency employees.[16] The Columbia Accident Investigation Board explicitly attributed the accident to

---

more variable than mean judgments of their members); Garold Stasser & William Titus, *Hidden Profiles: A Brief History*, 14 PSYCHOL. INQUIRY 304, 308–09 (2003) [hereinafter Stasser & Titus, *Hidden Profiles*] (discussing failures of groups to incorporate shared information).

[11] *See* IRVING L. JANIS, GROUPTHINK 7–9 (2d ed., rev. 1983) (discussing phenomenon of groupthink, by which people end up promoting consensus that does not reflect information held by individual group members).

[12] *See* REPORT OF THE 108TH CONGRESS, *supra* note 2, at 4–7.

[13] *Id.* at 4.

[14] *Id.* at 6.

[15] *Id.* at 8.

[16] 1 COLUMBIA ACCIDENT INVESTIGATION BD., NASA, THE COLUMBIA ACCIDENT INVESTIGATION BOARD REPORT 97–204 (2003), *available at* http://www.nasa.gov/columbia/home/CAIB_Vol1.html.

Imaged with Permission from N.Y.U. Law Review

NASA's unfortunate culture, one that does too little to elicit information. In the Board's words, NASA lacks "checks and balances."[17] It pressures people to follow a "party line."[18] At NASA, "it is difficult for minority and dissenting opinions to percolate up through the agency's hierarchy"[19]—even though, the Board contended, effective safety programs require the encouragement of minority opinions and a willingness to acknowledge, rather than to conceal, bad news.[20]

To explain why deliberation fails, I explore the consequences of two sets of influences on members of deliberating groups.[21] The first consists of informational influences, which cause group members to fail to disclose what they know because of deference to the information publicly announced by others. The second involves social pressures, which lead people to silence themselves in order to avoid reputational sanctions, such as the disapproval of peers and supervisors. As a result of these forces, groups often do not correct but instead amplify individual errors, emphasize shared information at the expense of unshared information, fall victim to cascade effects, and end up in a more extreme position in line with the predeliberation tendencies of their members.[22] Even federal judges are vulnerable to the relevant pressures, as both Republican and Democratic appointees show especially ideological voting when they are sitting with other judges appointed by presidents of the same political party.[23]

Because of those pressures, deliberative processes often fail to achieve their minimal goal of aggregating the information that the relevant deliberators actually have. Indeed, such processes often fail to aggregate information even as they decrease variance, and increase confidence, among their members.[24] A confident, cohesive, error-prone group is nothing to celebrate. On the contrary, it might be

---

[17] *Id.* at 12.

[18] *Id.* at 102 (internal citation omitted).

[19] *Id.* at 183.

[20] *Id.*

[21] I explore these mechanisms from a different direction in CASS R. SUNSTEIN, WHY SOCIETIES NEED DISSENT (2003), but without attention to statistical groups and information markets, and without focusing on amplification of errors, hidden profiles, and the common knowledge effect, which are major emphases here.

[22] This last possibility is emphasized in ROGER BROWN, SOCIAL PSYCHOLOGY: THE SECOND EDITION 200–45 (1986). *See also* SUNSTEIN, *supra* note 21, at 112.

[23] *See* Cass R. Sunstein, David Schkade & Lisa Michelle Ellman, *Ideological Voting on Federal Courts of Appeals: A Preliminary Investigation*, 90 VA. L. REV. 301, 304–06, 314 (2004) (showing effects of panel composition on judicial behavior).

[24] *See* Chip Heath & Rich Gonzalez, *Interaction with Others Increases Decision Confidence but Not Decision Quality: Evidence Against Information Collection Views of Interactive Decision Making*, 61 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 305, 306 (1995) (showing that deliberation does not necessarily increase accuracy, but it does increase confidence of group members).

Imaged with Permission from N.Y.U. Law Review

extremely dangerous both to itself and to others.[25] I shall suggest that these various problems raise serious conceptual and empirical questions about the accounts of deliberative democracy offered by Jürgen Habermas,[26] Fishkin,[27] Ackerman,[28] and Amy Gutmann and Dennis Thompson,[29] among others.[30] Suppose that deliberation is working in the way that theorists prescribe; suppose, that is, that participants are operating on terms of equality, nonstrategically, and otherwise in accordance with what might be seen as the internal morality of deliberative democracy.[31] Even then, deliberation might lead to extremely serious blunders for reasons that have yet to be addressed by theorists of deliberative democracy.[32] Another way to put the point is that deliberative democrats have failed to engage adequately with either cognitive or social psychology. An understanding of the widely dispersed nature of information and arguments, and of the difficulty of aggregating that information and those arguments through deliberative processes, provides an explanation of why serious blunders are likely to occur.

How might such blunders be avoided? As we shall see, information markets often outperform both statistical and deliberating groups, simply because they are so effective at pooling information. Indeed, information markets realign private incentives in a way that makes them exceptionally well-designed to reduce the problems that infect deliberating groups. Such markets are worth investigating in part because they provide an illuminating route by which to explore some characteristic defects in deliberative processes—and by which to identify reforms that should make them work better. If deliberative processes are to be improved, it might well be by building on the insights provided by the successes of information markets. The result of that rebuilding effort should be a revised and strengthened conception of the uses and limits of deliberative democracy, in a way that gives a better understanding of that idea—one that is more closely

---

[25] For a comparison of democratic and nondemocratic regimes concluding that open debate on policy makes regimes less vulnerable to overconfidence with respect to war, see DOMINIC D.P. JOHNSON, OVERCONFIDENCE AND WAR: THE HAVOC AND GLORY OF POSITIVE ILLUSIONS 180–83 (2004).

[26] *See generally* HABERMAS, *supra* note 4.

[27] *See generally* FISHKIN, *supra* note 5.

[28] *See generally* ACKERMAN & FISHKIN, *supra* note 6.

[29] *See generally* GUTMANN & THOMPSON, *supra* note 4.

[30] For a collection of views, see DELIBERATIVE DEMOCRACY, *supra* note 4.

[31] For a brief and clear statement of the preconditions for deliberation, see Jürgen Habermas, Between Facts and Norms: *An Author's Reflections*, 76 DENV. U. L. REV. 937, 940–41 (1999). For an overlapping view emphasizing moral requirements imposed on participants by the deliberative ideal, see GUTMANN & THOMPSON, *supra* note 4, at 39–51.

[32] *See infra* notes 85–102 and accompanying text.

Imaged with Permission from N.Y.U. Law Review

attuned to both cognitive and social psychology. In addition, information markets are worth investigating in their own right, if only because they promise to provide a supplement to deliberation, one that should improve social decisions.[33]

One of my chief goals, in short, is to mend deliberative processes, not to end them. As we shall see, both social norms and institutional design can go a long way toward reducing the problems that lead to deliberative blunders in a way that has implications for the performance of many institutions, including legislatures, committees, and even multimember judicial panels.

To keep the analysis simple, I shall focus not on controversial judgments of value but on questions with demonstrably correct answers. An understanding of how deliberation finds, and fails to find, those answers should have implications for its potential and its limitations with respect to normative questions as well. If deliberation often fails to produce good answers to simple questions of fact, then it is also likely to fail to produce good answers to disputed issues of value. The solution to many such questions depends at least in part on resolution of factual issues; it is difficult to take a stand on proposals to raise the minimum wage, to engage in preemptive war, or to overrule *Roe v. Wade*[34] without resolving several issues of fact. And even when factual issues are not central, deliberation can, in principle, ensure more sensible judgments.[35] Unfortunately, however, the problems posed by informational pressure and social influences apply in normative domains not less than elsewhere. I will therefore offer some suggestions for how groups can reduce those problems through structural reforms.

This Article proceeds in five parts. Part I begins the analysis with a description of a simple, nondeliberative method for aggregating privately-held information, one that takes the average of predeliberation judgments. The resulting judgments of these "statistical groups" are sometimes remarkably accurate. They also provide a useful benchmark for assessing deliberative judgments.[36] A key goal of deliberation is to improve on the predeliberation judgments of individuals and groups; if we understand how statistical groups perform, we can have

---

[33] For an ambitious account, see Robin Hanson, Shall We Vote on Values, but Bet on Beliefs? 5–27 (Sept. 2003) (unpublished manuscript, on file with the *New York University Law Review*).

[34] 410 U.S. 113 (1973).

[35] This is the thesis of FISHKIN, *supra* note 5, at 16, 40–43.

[36] These often are described as the judgments of "statisticized groups." *See* Irving Lorge et al., *A Survey of Studies Contrasting the Quality of Group Performance and Individual Performance, 1920–1957*, 55 PSYCHOL. BULL. 337, 344 (1958).

Imaged with Permission from N.Y.U. Law Review

a better sense of the potential virtues of deliberation and also a better appreciation of what discussion might be expected to accomplish.

Part II explores how deliberation might be expected, in principle, to improve on the judgments of statistical groups. Unfortunately, the expectation of improvement is often dashed, largely because of the effects of informational pressures and social influences. These pressures and influences contribute to the amplification of errors, hidden profiles, cascade effects, and group polarization. Part II explores these problems and attempts to square some apparently conflicting evidence about the performance of deliberating groups; it investigates the possibility that some groups will do as well as or even better than their best members.

Part III outlines structural reforms that are intended to ensure that group members reveal what they know—for example, by requiring anonymous statements of beliefs before deliberation begins, by assigning specified roles to participants in deliberation, and by structuring incentives to produce disclosure of privately held information. Two reforms are especially important here: "priming" participants by asking them to think critically rather than to behave cooperatively and asking people to attempt to improve the group's overall decision, rather than to make an individually correct choice.

Part IV identifies and compares information markets, in which people bet on the outcomes of events. Information markets have performed remarkably well in many diverse settings—by, for example, calling the 2004 presidential election with uncanny accuracy[37] and also predicting all eight of the major Oscar winners in 2005.[38] Because information markets restructure people's incentives, overcome collective action problems faced by individual group members, and allow informed traders to play a large role in setting "prices," they have advantages over both statistical judgments and deliberative judgments. They might well be used as a supplement to or even a replacement for collective deliberation. Part V briefly discusses how the analysis might apply to normative questions.

I
STATISTICAL GROUPS

Suppose that there is a question about some disputed issue of fact. The issue might involve past events. How many home runs did Hank Aaron hit? When was Calvin Coolidge elected president? Or

[37] *See infra* notes 71–73 and accompanying text.

[38] I am referring here to the Hollywood Stock Exchange. *See infra* notes 333–36 and accompanying text.

Imaged with Permission from N.Y.U. Law Review

the issue might involve a prediction about the future. Will a district court decision be reversed on appeal? Does a foreign country pose a serious threat to national security? Is the United States likely to have difficulty in winning a particular war? Will the poverty rate, or concentrations of specified pollutants, increase in the next year? A great deal of evidence suggests that, under certain conditions, a promising way to answer such questions is this: *Ask a large number of people and take the mean answer.*[39] When the relevant conditions are met,[40] the mean answer, which we might describe as the group's "statistical answer," is often accurate, where accuracy is measured by reference to objectively demonstrable facts.

Consider a simple example. In 2004, members of the Society for American Baseball Research were asked, in advance, to predict the winners of the baseball playoffs.[41] Remarkably, strong majorities of the 413 respondents correctly predicted all of the first round winners: New York, Boston, Houston, and St. Louis. At least as remarkably, a majority predicted, correctly, that St. Louis would win the National League pennant, and a large plurality predicted that the Red Sox would win the American League pennant. A plurality also favored the Red Sox to win the World Series. Hence the favored choice of the expert group was right 100% of the time.

If this striking finding can be generalized, it is a helpful way to approach the uses of both deliberation and information markets. If statistical answers are often accurate, then accuracy on the part of deliberating groups might not be a product of deliberation at all. Perhaps deliberation does not much matter, and accuracy simply comes from the fact that significant numbers of people are being asked to state their views. And if statistical answers are accurate, information markets might be supremely accurate, simply because investors are being asked to back their views with dollars. Consider here the fact that with statistical groups, individual judgments are in a sense "pure"; they are independent and hence unaffected by an understanding of what other group members think. So too with information markets, in which there is no public give-and-take among group members.[42] Are

---

[39] When groups are large, of course, and not subject to systematic biases, the mean and the median will tend to converge.

[40] *See infra* notes 42–50 and accompanying text.

[41] John Zajc, THIS WEEK IN SABR (Soc'y for Am. Baseball Research, Cleveland, Ohio), Oct. 9, 2004 (Results of playoff prediction survey), *at* http://www.sabr.org/sabr.cfm?a=cms,c,1123,3,212.

[42] Of course investment decisions may result from social deliberation, and the individual judgments that enter into the judgments of statistical groups are often affected by such deliberation. The point is that group deliberation is, by definition, not a part of statistical aggregation through either statistical groups or the price mechanism.

Imaged with Permission from N.Y.U. Law Review

statistical groups, or information markets, more or less likely to be accurate because of the absence of deliberation?

It is well-known that statistical answers from groups of sufficiently large sizes tend to match the views of population-wide samples.[43] This finding bears on issues as diverse as the use of juries as a measure of community sentiment[44] and the remarkable success of Google, the search engine; Google is good at finding what a particular searcher wants because it knows what most searchers want.[45] But here the question is what is true, not what populations think. Let us therefore explore how statistical groups actually perform.

A. Evidence of Statistical Group Accuracy

Many of the studies of statistical groups involve quantitative estimates. Consider a few examples:

1. In an early study, Hazel Knight asked college students to estimate the temperature of a classroom.[46] Individual judgments ranged from 60 degrees to 85 degrees; the statistical judgment of the group was 72.4 degrees, very close to the actual temperature of 72 degrees. That judgment was better than 80% of the individual judgments.
2. Judging the numbers of beans in a jar, the group average is almost always better than the judgments of the vast majority of individual members. In one such experiment, a group of fifty-six students was asked about a jar containing 850 beans; the group estimate was 871, a better guess than all but one of the students.[47]
3. Asking two hundred students to rank items by weight, one experimenter found that the group's average estimate was 94% accurate—a figure surpassed by only five individuals.[48]

---

[43] *See* H.J. Eysenck, *The Validity of Judgments as a Function of Number of Judges*, 25 J. EXPERIMENTAL PSYCHOL. 650, 651 (1939) (showing that large groups reflect judgments of larger populations).

[44] *See* Cass R. Sunstein et al., *Assessing Punitive Damages*, 107 YALE L.J. 2071, 2095–99 (1998) (showing that small groups often reflect judgments of community as whole, at least when their judgments are made on bounded scale).

[45] *See* Sergey Brin & Lawrence Page, *The Anatomy of a Large-Scale Hypertextual Web Search Engine*, 30 COMPUTER NETWORKS & ISDN SYS. 107, 107–10 (1998), *available at* http://newdbpubs.stanford.edu:8090/pub/1998-8.

[46] Lorge et al., *supra* note 36, at 344.

[47] *See* JAMES SUROWIECKI, THE WISDOM OF CROWDS 5 (2004) (discussing jar experiment).

[48] *Id.*; Kate Gordon, *Further Observations on Group Judgments of Lifted Weights*, 1 J. PSYCHOL. 105, 106 (1936); Kate Gordon, *Group Judgments in the Field of Lifted Weights*, 7

Imaged with Permission from N.Y.U. Law Review

4. Asked to rank ten piles of buckshot, each only slightly different in size from the others, the combined group's average guess was 94.5% accurate, far more so than that of almost all individual group members.[49]

5. The British scientist Francis Galton sought to draw lessons about collective intelligence by examining a competition in which contestants attempted to judge the weight of a fat ox at a regional fair in England. The ox weighed 1198 pounds; the average guess, from the 787 contestants, was 1197 pounds.[50]

In light of these findings, many questions might plausibly be answered not deliberatively, but simply by asking a large group of people and selecting the average response. Imagine that a large company is attempting to project its sales for the following year. Might it do best to poll its salespeople and to choose the average number on the assumption that it is likely to be correct?[51] Or suppose that a company is deciding whether to hire a new employee. Should it ask relevant personnel, not to deliberate, but about their individual views on whether the employee's performance is likely to meet a certain level? Or turn to the legal context and suppose that the question is whether a case should be settled. Ought a law firm to poll its lawyers about the expected outcome at trial? Or consider the political domain and suppose that the question is whether a war effort, or an environmental initiative, will go well by some identifiable standard. Should the President poll his advisers and take the median answer? To answer these questions, we have to know why, in the relevant studies, the median judgment is so accurate.

B. The Condorcet Jury Theorem

The accuracy of judgments of statistical groups is best explained by reference to the Condorcet Jury Theorem.[52] To see how the Jury Theorem works, suppose that people are answering the same question with two possible answers, one false and one true, and that the average probability that each voter will answer correctly exceeds 50%.

J. EXPERIMENTAL PSYCHOL. 398, 399–400 (1924) (offering evidence on assessments of weights).

[49] Richard S. Bruce, *Group Judgments in the Field of Lifted Weights and Visual Discrimination*, 1 J. PSYCHOL. 117, 117–21 (1936).

[50] SUROWIECKI, *supra* note 47, at xi–xiii.

[51] Some affirmative evidence can be found in Armstrong, *supra* note 3, at 417, 419–20, 427, 433–35.

[52] *See* William P. Bottom et al., *Propagation of Individual Bias Through Group Judgment: Error in the Treatment of Asymmetrically Informative Signals*, 25 J. RISK & UNCERTAINTY 147, 152–54 (2002) (describing Condorcet Jury Theorem).

Imaged with Permission from N.Y.U. Law Review

The Jury Theorem holds that the probability of a correct answer, by a majority of the group, increases toward certainty as the size of the group increases.[53] The importance of the Jury Theorem lies in the demonstration that groups are likely to do better than individuals, and large groups better than small ones, if majority rule is used and if each person is more likely than not to be correct. The last proviso is extremely important. Suppose that each individual in a group is more likely to be wrong than right. If so, the likelihood that the group will decide correctly falls to zero as the size of the group increases.

In the context of statistical judgments, several of Condorcet's stringent and somewhat unrealistic assumptions are met. Indeed, the likelihood that they will be met is higher with statistical groups than with deliberating ones. Condorcet assumed that people would be unaffected by whether their votes would be decisive, that people would not be affected by one another's votes, and that the probability that one group member would be right would be statistically unrelated to the probability that another group member would be right.[54] The first two assumptions plainly hold for statistical groups. People do not know what others are saying and hence they cannot be influenced by a belief that their judgments will make the difference to that of the group. The third assumption may or may not be violated. Those who have similar training, or who work closely together, will be likely to see things in the same way,[55] and those involved in statistical groups might well meet these conditions. On the other hand, the Condorcet Jury Theorem has been shown to be robust to violations of this third assumption.[56]

To see why statistical groups perform well, consider the problems just described and note that even if everyone in the group is not more than 50% likely to be right, the Theorem's predictions may well continue to hold. Suppose, for example, that 60% of people are 51% likely to be right and that 40% of people are 50% likely to be right; or that 45% of people are 40% likely to be right and that 55% of people are 65% likely to be right; or even that 51% of people are 51% likely to be right and that 49% of people are merely 50% likely to be right. Even under these conditions, the likelihood of a correct answer will move toward 100% as the size of the group increases. It will not move as quickly as it would if every group member were highly likely to be

---

[53] The theorem is based on some simple arithmetic. Suppose, for example, that there is a three-person group, in which each member has a 67% probability of being right. The probability that a majority vote will produce the correct answer is 74%. *Cf. id.* at 153.

[54] *See id.* at 153.

[55] *Id.*

[56] *Id.*

Imaged with Permission from N.Y.U. Law Review

right, but it will nonetheless move. We could imagine endless variations on these numbers. The point is that even if a significant percentage of the group is not more likely to be right than wrong, or even if many group members are more likely to be wrong than right, an accurate result can be expected if the group is sufficiently large.

Of course most of the relevant judgments in studies of statistical groups do not involve a binary choice; consider the question how many beans are in a jar, how many pounds a given object weighs, or how many copies of a certain book will sell in the following year. But the answers to such questions are not analytically different from those in binary choices. In answering the relevant questions, each person is effectively being asked to answer a long series of binary questions—ten beans or a thousand beans, twenty beans or five hundred beans, fifty beans or one hundred beans, and so on. If a sufficiently large group is asked to answer such questions, and if most individual answers will be better than random, the mean answer will be highly accurate. Of course the combination of probabilities for a series of binary results might mean that things will turn out poorly. If someone is 51% likely to answer *each* of two questions correctly, the probability that she will answer *both* questions correctly is only slightly higher than 25%. But with large groups, enough people are likely to make better-than-random guesses on the questions involved in certain quantitative judgments that the average estimate will have a high degree of accuracy.[57]

Compare a situation in which only 49% of the group is likely to be better than random. If so, the likelihood of a mistake will move toward 100% under the same condition. But for the number of beans in a jar, or the weight of an ox, most people are not wholly at sea. The accuracy of the median judgment, for large groups, is simply an application of the Condorcet Jury Theorem. And, in certain circumstances, deliberating groups will act in roughly the same way, aggregating their information to produce remarkably accurate results.[58]

---

[57] On some of the technical complexities, see Christian List & Robert E. Goodin, *Epistemic Democracy: Generalizing the Condorcet Jury Theorem*, 9 J. POL. PHIL. 277, 283–88, 295–97 (2001). For a popular illustration, consider the television show *Who Wants to Be a Millionaire?* In this show, contestants, when stumped, are permitted to ask a personally appointed "expert" (a friend who the contestant thinks knows a great deal) or the studio audience. The studio audience significantly outperforms the expert. *See* SUROWIECKI, *supra* note 47, at 3–4.

[58] Bottom et al., *supra* note 52, at 160–61.

Imaged with Permission from N.Y.U. Law Review

C. Errors

In this light, we can identify two sets of situations in which the judgment of a statistical group will be incorrect. The first are those in which group members show a systematic bias. The second, a generalization of the first, are those in which their answers are worse than random. The failures of statistical judgments in these circumstances have strong implications for deliberation as well.

1. Bias

A systematic bias in one or another direction will create serious problems for the group's answers. If, for example, an experimenter "anchors" subjects on a misleading number, the median will almost certainly be wrong. Suppose that a jar contains 800 jelly beans, and the experimenter happens to say, quietly, "Many jars of jelly beans, though not necessarily this one, have 500 jelly beans," or even, "I'm asking this question to 250 people."[59] In either case, the low number will likely operate as an anchor,[60] and people's answers will be systematically biased toward understating the actual number, producing an unreliable median. One study demonstrates more generally that a group's statistical estimate is likely to be erroneous "when the material is unfamiliar, distorted in a way such that all individuals are prone to make similar errors of estimation."[61]

Anchors have significant effects within the legal system. For example, the plaintiff's demand is likely to affect damage awards for harms that are difficult to monetize. Groups are no less subject to those effects than individuals.[62] Even judges have been found to be subject to irrelevant anchors,[63] and there is every reason to believe that multimember courts would be at least as vulnerable to them as individual judges are.[64]

---

[59] Even self-evidently arbitrary anchors have significant effects on people's judgments. *See* Gretchen Chapman & Eric Johnson, *Incorporating the Irrelevant: Anchors in Judgments of Belief and Value*, *in* HEURISTICS & BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT 120 (Thomas Gilovich et al. eds., 2002).

[60] *See id.*

[61] Lorge et al., *supra* note 36, at 346.

[62] *See* Reid Hastie et al., *Do Plaintiffs' Requests and Plaintiffs' Identities Matter?*, *in* CASS R. SUNSTEIN ET AL., PUNITIVE DAMAGES: HOW JURIES DECIDE 62, 73–74 (2002). *See generally* Chapman & Johnson, *supra* note 59.

[63] Chris Guthrie et al., *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 790–91 (2001) (showing effect of anchors on judicial judgments in experimental contexts).

[64] *See* Norbert L. Kerr et al., *Bias in Judgment: Comparing Individuals and Groups*, 103 PSYCHOL. REV. 687, 689, 691–93 (1996) (noting studies showing that anchors affect groups as well as individuals).

Imaged with Permission from N.Y.U. Law Review

2. *Random or Worse*

Suppose that people are asked not about the number of jelly beans in a jar, but about the number of atoms in a jelly bean. On that question, most people's answers are hopelessly ill-informed, and there is no reason at all to trust their judgments. Consider a small-scale study at the University of Chicago Law School, one that strongly supports this conclusion. A number of faculty members were asked the weight, in pounds, of the fuel that powers space shuttles.[65] The actual answer is four million pounds. The median response was 200,000; the mean was 55,790,555 (because of one outlier choice)—both wildly inaccurate. In a binary choice, of course, people's answers will be worse than random only if they are unaware of how little they know; if they know that they are likely to be wrong, they should choose randomly, which gives them a 50% probability of being right. But sometimes people think they know more than they do, and many tasks do not involve binary choices at all. Statistical groups will err if confusion and ignorance are so widespread that individuals' answers are worse than random.

Here, too, there are evident applications to many contexts in law and politics. If, for example, members of Congress make systematic blunders about the risks caused by certain pollutants, or if federal judges systematically err about the competitive effects of certain practices said to violate the antitrust laws, then legislative and judicial judgments will be erroneous even though the beliefs of numerous people are being taken into account.

D. Statistical Answers and Experts

Should statistical means be used more frequently than they now are? Do statistical means outperform experts? Everything depends on the competence of the experts. If we could find real experts on estimating the weight of oxen or on counting jelly beans, and if we understand expertise to be the ability to make accurate assessments, then experts would, by definition, do better than statistical means. Suppose, for example, that a deliberating group of lawyers is trying to decide how many Supreme Court decisions have invalidated a state or federal law, or the number of lines in *Antigone*, or the weight of the most recent winner of the Kentucky Derby. Would it make any sense to poll the lawyers individually and to assume that the mean response is accurate? The studies outlined above suggest that if the group is large enough, the mean answer will be quite good, at least if group

[65] Unpublished data, available from author (on file with the *New York University Law Review*).

Imaged with Permission from N.Y.U. Law Review

members are not systematically biased and if they are more likely than not to be right.[66] But there are many ways to do far better.

If experts are available, it would make sense to obtain a statistical answer from a group of them, rather than to select one or a few. If experts are likely to be right, a statistical group of experts should have the same advantage over individual experts as a statistical group of ordinary people has over ordinary individuals. In fact a great deal of evidence supports this claim.[67] Return to the 2004 baseball predictions described above; in that context, the judgments of statistical groups were uncannily accurate. Because those judgments were perfect as a group, they could not possibly be inferior to that of any individual expert.

There is more systematic evidence in this vein. In a series of thirty comparisons, statistical groups of experts had 12.5% fewer errors than individual experts on forecasting tasks involving such diverse issues as company earnings, cattle and chicken prices, real and nominal GNP, survival of patients, and housing starts.[68] For example, statistical groups of experts significantly outperformed individual experts in predicting the annual earnings of firms, changes in the American economy, and annual peak rainfall runoff.[69] The implication is straightforward: "Organizations often call on the best expert they can find to make important forecasts. They should avoid this practice, and instead combine forecasts from a number of experts."[70] For political polling, it has become standard practice to combine a set of poll results and to rely on the mean or median, rather than to select one or two.[71] The most sophisticated treatment here involves "Polly," a program designed to predict the results of the 2004 presidential election.[72] Polly made her predictions after combining a large set of sources: polls, computer models, expert panels, and information mar-

---

[66] I conducted such a poll with faculty at the University of Chicago Law School, who did fairly well in estimating the weight of the horse who won the Kentucky Derby, fairly badly in estimating the number of lines in *Antigone*—and horrendously with the number of Supreme Court invalidations of state and federal law! (Author's notes on file with the *New York University Law Review*).

[67] *See* Armstrong, *supra* note 3, at 419–20. For many factual questions, of course, a little research would be sufficient to identify the correct answers. But for some factual issues, even significant research is inconclusive, and it is best to consult experts.

[68] *Id.* at 428.

[69] *Id.* at 428, 430–31.

[70] *Id.* at 433.

[71] *See, e.g.*, Sam Wang, Electoral College Meta-Analysis (2004) (combining predictions and discussing rationale for doing so), *at* http://synapse.princeton.edu/~sam/pollcalc.html (last visited Feb. 9, 2005).

[72] *See* Alfred Cuzán et al., Combining Methods to Forecast the 2004 Presidential Election: The Pollyvote 12 (Jan. 6, 2005) (unpublished manuscript, on file with the *New York University Law Review*), *available at* http://www.politicalforecasting.com.

Imaged with Permission from N.Y.U. Law Review

kets. Over eight months, she continued to predict that President Bush would receive at least 50% of the vote; her final forecast was that President Bush would receive 51.5% of the vote, a number that was very close to precisely right.[73]

Consider in this regard the Copenhagen Consensus, generated by a group of economists in an attempt to inform policy judgments about global risks.[74] The Copenhagen Consensus emerged from an effort to explore a series of possible interventions, involving climate change, water and sanitation, hunger and malnutrition, free trade, and communicable diseases, among others. A number of experts were asked about the best way to promote global welfare, and particularly the welfare of developing countries, assuming that $50 billion were made available. The experts ranked the possible projects, producing an overall ranking (reflecting the mean rankings of the experts taken as a whole).[75] I do not mean to suggest that the results of this particular exercise are correct; everything depends on whether the relevant experts were in a position to offer reliable answers on the questions at hand. But if statistical means are a good way to aggregate knowledge when ordinary people know something of relevance, then they are also a good way to aggregate knowledge from experts.[76]

II
DELIBERATING GROUPS

Although the judgments of statistical groups can be quite accurate, it is easy to imagine that a deliberating group would be much better. In principle, a deliberating group should do well even when its members are error-prone. Deliberation, in the form of an exchange of information and reasons, might well bring them into line. If many group members give answers that are worse than random, perhaps

---

[73] *See id.* at 1–2, 12. Polly's final forecast was that President Bush would take 51.5% of the two-party vote, which was exceedingly close to the final tally. *Id.* at 12; *see also* Political Forecasting, Polly's Page, *at* www.politicalforecasting.com (last updated Jan. 5, 2005) ("POLLY WAS RIGHT! Who would win in November? . . . This question consumed Polly since March, when her page was launched. She heard from many sources, including 268 polls, 10 forecasting models, three surveys . . . of a select panel of American politics experts, and the Iowa Electronic Markets (IEM).").

[74] *See* GLOBAL CRISES, GLOBAL SOLUTIONS 1–9 (Bjørn Lomborg ed., 2004).

[75] *See id.* at 605–08.

[76] In the context of the Copenhagen Consensus, the assessment of the central question—"the best ways of advancing global welfare"—involved normative judgments as well as factual ones. For example, the experts ranked control of malaria above improving infant and child nutrition, *see id.* at 606, but this ranking involves contentious judgments of value as well as empirical claims. The simplest way to test the aggregation of expert judgments is through use of prediction questions on which unambiguous evidence is available. On normative questions, see *infra* Part V.

Imaged with Permission from N.Y.U. Law Review

other group members can show them how they have erred. If individuals have been manipulated in their private judgments, perhaps deliberation will undo the effects of the manipulation.[77] If individual members have anchored on a misleading value, perhaps deliberation will expose the anchor as such.

History attests to the widespread belief that deliberation is likely to improve individual judgments. In countless domains, both public and private, deliberation is the preferred method for arriving at the correct answer, even or perhaps especially when the stakes are high. Certainly this has long been true in legislatures and on corporate boards; it is also true for multimember panels and in academia. No one doubts that it is possible to find numerous contexts in which deliberative processes appear to have done a great deal of good. The question is whether some such processes do better than others, and whether it is really sensible to celebrate deliberation as such and in the abstract. I will be raising a number of doubts about such celebrations.

To make the analysis tractable, let us focus on how deliberating groups might be able to solve factual questions or cognitive puzzles that have correct solutions. These questions and puzzles often involve instrumental rationality, asking people to identify the right strategy for achieving agreed-upon goals. Results in these domains provide a good test of when and whether deliberating groups perform well. To the extent that such groups do badly in answering questions with objectively correct answers, we have reason to suspect that they will also do badly in answering questions for which there is no consensus on truth or validity.

A. Mechanisms and Realities

1. Possibilities

Because of the Condorcet Jury Theorem, we might be tempted to think that deliberating groups will perform extremely well. Note, however, that one of Condorcet's key assumptions does not hold for such groups: Members are not making their judgments independently. Perhaps this is a virtue rather than a vice. If groups perform better than their average member, we can imagine three principal mechanisms by which the improvement occurs:

(1) Groups might operate in such a way as to equal the performance of their best members. One or more group members will often know the right answer, and other group members might well become

[77] *See* ACKERMAN & FISHKIN, *supra* note 6, at 52–59 (emphasizing desirable effects of deliberation).

Imaged with Permission from N.Y.U. Law Review

convinced of this fact. For this reason, groups might perform toward or at the level of their best members. If some or many members suffer from ignorance, or from a form of bias that leads to error, other group members might correct them. Suppose, for example, that a panel of judges is trying to recall relevant Supreme Court decisions in a somewhat specialized area. If one of the judges is actually aware of those decisions, the group will be made aware of them too. Or suppose that a group of military officials is attempting to assess the strengths and weaknesses of a potential enemy in some part of the world. If one of them is a specialist, all of them can learn what the specialist knows. Many deliberating groups contain at least one expert on the question at hand; if group members listen to the expert, they will do at least as well as she does. For these reasons, deliberation might correct individual errors, rather than propagating them, in a way that allows convergence on the judgment of the most accurate group member.

(2) Deliberation could aggregate existing information in a way that leads the group as a whole to know more than any individual member does. Suppose that the group contains no experts on the question at issue, but that relevant information is dispersed among group members so that the group is potentially expert even if its members are not. Or suppose that the group contains a number of experts, but that each member is puzzled about how to solve a particular problem. Deliberation might elicit the relevant information and allow the group to make a sensible judgment. Almost everyone has had the experience of being a part of a group that ended up with a solution that went beyond what any individual member could have produced on her own. In this process, the whole is equal to the sum of the parts—and the sum of the parts is what is sought.

(3) The give-and-take of group discussion might sift information and perspectives in a way that leads the group to a good solution to a problem, one in which the whole is actually *more* than the sum of its parts. In such cases, deliberation is, at the very least, an ambitious form of information aggregation, one in which the exchange of views leads to a creative answer or solution. And in fact, groups sometimes do outperform their best members.[78]

2. *Variance, Confidence, and Legitimacy*

To what extent do these mechanisms work in practice? Two points are entirely clear. First, deliberation usually reduces variance.[79]

[78] *See* Gigone & Hastie, *Proper Analysis*, *supra* note 1, at 143–53 (offering some examples of group success, while showing that such success is not typical).

[79] *See* BROWN, *supra* note 22, at 206–07 (showing reduction in variance).

Imaged with Permission from N.Y.U. Law Review

After talking together, group members tend to come into accord with one another.[80] Statistical groups thus show far more diversity of opinion than deliberating groups. Second, group members tend to become far more confident of their judgments after they speak with one another.[81] A significant effect of group interaction is a greater sense that one's post-deliberation conclusion is correct—whether or not it actually is.[82] Corroboration by others increases confidence in one's judgments.[83] It follows that members of deliberating groups will usually converge on a position on which group members have a great deal of confidence. This is not disturbing if that position is also likely to be correct—but if it is not, then many group members will end up sharing a view in which they firmly believe, but which turns out to be wrong (a most unfortunate and sometimes quite dangerous situation[84]).

If the purpose of deliberation is not simply to produce accurate outcomes, then it might be especially important to know that deliberation ensures less variance and higher confidence. Suppose that a key goal of deliberation is to promote a sense of *legitimacy*—an understanding, by group members, that they have been able to participate in the process and a belief, by all concerned, that the decision is acceptable on its merits. Because deliberation decreases variance and increases confidence in the outcome, it might be favored even if it produces errors. Decreased variance and increased confidence may well be significant and independent goods.

In fact there are complex tradeoffs among the relevant variables here. Should we be willing to accept a little more error if that is the price for less variance and more confidence? If deliberation significantly increases confidence and a sense of legitimacy, then it might be desirable even if the decision is slightly worse—at least if little turns on slight differences in the quality of the outcome. Perhaps what most matters is that group members accept the decision, not that the decision be correct. At least this is so if the stakes are relatively low—if it does not greatly matter whether the group chooses course A or course

---

[80] *Id.*

[81] *See generally* Heath & Gonzalez, *supra* note 24 (showing increase in confidence without increase in accuracy).

[82] *See* RUPERT BROWN, GROUP PROCESSES 176 (2d ed. 2000) ("The groups were also more confident about the correctness of their answers, and this was true even when they got the answers wrong!").

[83] *See* Robert S. Baron et al., *Social Corroboration and Opinion Extremity*, 32 J. EXPERIMENTAL SOC. PSYCHOL. 537, 538 (1996) (discussing effects of corroboration in increasing extremism).

[84] *See* JOHNSON, *supra* note 25, at 181–86 (showing that overconfidence often leads to war-making, with disastrous results).

Imaged with Permission from N.Y.U. Law Review

B. On the other hand, an increase in legitimacy might not be so important if the decision is leading the group into a serious blunder. If course A is really a great deal better than course B, it is probably worthwhile to ensure that deliberation will produce the superior result. For many decisions, a key goal of deliberation is to improve choices, not to legitimate whatever choice ultimately is made. And if deliberation can improve outcomes while also increasing legitimacy, so much the better.

3. *Accuracy: A Mixed Verdict*

Unfortunately, there is no systematic evidence that deliberating groups will usually succeed in aggregating the information that their members have.[85] This finding presents an extremely serious problem for those who favor deliberation as a method for improving judgments. My ultimate goal is to use this evidence to explain how deliberative processes should be improved; hence the goal is not to eliminate or to reduce deliberation, but to use the mixed findings to identify institutional reforms that will increase accuracy. Let us see what the evidence shows.

With respect to questions with definite answers, deliberating groups tend to do about as well as or slightly better than their average member, but not as well as their best members.[86] Hence it is false to say that group members usually end up deferring to their internal specialists. Truth does not win out; the most that can be said is that, under some conditions, the group will converge on the truth if the truth begins with "at least some initial social support" within the group when the task has "a demonstrably correct answer according to a broadly shared normative framework (e.g., deductive logic)."[87] Note here that when groups outperform most of their individual mem-

---

[85] *See* BROWN, *supra* note 82, at 173–93 (describing complex results about group performance); Gigone & Hastie, *Proper Analysis*, *supra* note 1, at 149, 153 (summarizing studies as finding group judgments to be approximately equal to accuracy of mean judgments of their members, and less accurate than judgments of their most accurate member); Kerr et al., *supra* note 64, at 713 (finding, after canvassing of empirical literature, "no simple empirical answer" to question of whether individuals or groups are more likely to make biased judgments).

[86] *See* Gigone & Hastie, *Proper Analysis*, *supra* note 1, at 153 (summarizing findings that groups do not perform as well as best members); Reid Hastie, Review Essay: *Experimental Evidence of Group Accuracy, in* INFORMATION POOLING AND GROUP DECISION MAKING 129, 133–46 (Bernard Grofman & Guillermo Owen eds., 1983) (same). To the same effect, see Garold Stasser & Beth Dietz-Uhler, *Collective Choice, Judgment, and Problem Solving, in* BLACKWELL HANDBOOK OF GROUP PSYCHOLOGY: GROUP PROCESSES 31, 49–50 (Michael A. Hogg & R. Scott Tindale eds., 2001) [hereinafter GROUP PROCESSES] (collecting findings).

[87] MacCoun, *supra* note 10, at 120.

Imaged with Permission from N.Y.U. Law Review

bers, it is generally because the issue is one on which a particular answer can be shown, to the satisfaction of all or most, to be right; and that even in that condition, the group might not do well if the demonstrably correct solution lacks significant support at the outset.[88] On brainteasers and crossword puzzles, on the other hand, groups tend to perform better than individuals, because they engage in a process of information aggregation and mutual error correction.[89] But with respect to brainstorming problems, deliberating groups have been found to do far less well than statistical groups, apparantly because deliberating groups discourage novelty.[90] Hence "brainstorming is actually most beneficial when carried out initially in *private*, the interacting group then being used as a forum for combining and evaluating these individually produced ideas."[91]

No significant differences are found between deliberating groups and *average* individual performances in numerical estimates, such as assessment of the number of beans in a jar or the length of lines.[92] One study finds that when asked to estimate the populations of American cities, groups did as well as their most accurate individual member;[93] but this is an atypical result.[94] Another study attempted to test whether deliberating groups were particularly good at telling whether people were telling the truth or lying.[95] The individual votes, predeliberation, were 48% correct, about the same as the post-deliberation judgments. Approximately the same number of people shifted toward error as toward correct answers. Yet another study finds that in various brainteasers, groups were better than their average member, but not as good as their best member.[96] Still other studies find that in estimating quantities, groups do about as well as their average member, and not as well as their best member.[97]

In general, simple majority schemes do fairly well at predicting group judgments for many decisionmaking tasks. It follows that if the

---

88 *See id.*

89 *See* Stasser & Dietz-Uhler, *supra* note 86, at 35.

90 *See* BROWN, *supra* note 82, at 176.

91 *Id.* (emphasis in original).

92 *See* Hastie, *supra* note 86, at 133.

93 Hillel J. Einhorn et al., *Quality of Group Judgment*, 84 PSYCHOL. BULL. 158, 168 (1977).

94 *See* Hastie, *supra* note 86, at 133–46 (showing that groups do not usually perform as well as their most accurate individual member).

95 *See* Harold E. Burtt, *Sex Differences in the Effect of Discussion*, 3 J. EXPERIMENTAL PSYCHOL. 390 at 390–95 (1920).

96 *See* Hastie, *supra* note 86, at 147.

97 *See id.* at 134–38.

Imaged with Permission from N.Y.U. Law Review

majority is wrong, the group will be wrong as well.[98] With experts, the same general conclusion holds. A "structured approach for combining independent forecasts is invariably more accurate" than "traditional group meetings," which do "not use information efficiently."[99] These points raise serious doubts about the celebrations of deliberative democracy offered by Habermas[100] and others,[101] simply because those celebrations do not engage with the empirical findings, and because the preconditions for deliberation do not provide an adequate safeguard against errors (as we shall see).[102]

Let us now turn to the key sources of deliberative failure, understood as a failure to make accurate decisions on the basis of the information that group members actually have.

B. Two Sources of Deliberative Failure: Informational Influences and Social Pressures

A primary advantage of statistical groups is that members say what they think. But with deliberating groups, this might not happen. Exposure to the views of others might lead people to silence themselves, and for two different reasons.

1. Information

The first reason involves the informational signals provided by the acts and views of other people. If most group members believe that X is true, there is reason to believe that X is in fact true. That reason might outweigh the purely private reason that a particular group member has to believe that X is false, and hence that particular member might simply defer. If most group members share a particular belief, isolated members, or members with a minority view, might not speak out, respecting the informational signal given by the statements of others.[103]

Not surprisingly, the strength of the signal will depend on the number and nature of the people who are giving it. People are particularly averse to being the sole dissenter. If all but one person in a deliberating group have said that X is true, then the remaining

---

98 *See* MacCoun, *supra* note 10, at 124 (showing that individual biases are often amplified by group interaction, not diminished).

99 Armstrong, *supra* note 3, at 433.

100 *See* HABERMAS, *supra* note 4, at 287–328.

101 *See* ACKERMAN & FISHKIN, *supra* note 6, at 1–16; GUTMANN & THOMPSON, *supra* note 4, at 1–9.

102 *See infra* Part II.B.5 (discussing Hayekian challenge to Habermas).

103 *Cf.* Andrew Caplin & John Leahy, *Miracle on Sixth Avenue: Information Externalities and Search*, 108 ECON. J. 60, 60–61, 63 (1998) (discussing information externalities created by behavior of other actors).

Imaged with Permission from N.Y.U. Law Review

member is likely to agree that X is true, even to the point of ignoring the evidence of his or her own senses.[104] And if the group contains one or more people who are well known to be authorities, then other group members are likely to defer to them.[105]

Informational signals come in three different forms involving conduct, conclusions, and reason-giving. First, group members might purchase certain products, visit particular places, or engage in certain actions; their conduct will provide a signal about their beliefs. Second, group members might express their conclusions about some issue. They might say that global warming is a serious problem, that crime is rising in New York City, or that minimum wage legislation increases unemployment. Third, group members might give reasons and arguments for their beliefs, going beyond conclusions to explain why they think as they do. If a number of different arguments favor a certain conclusion, and if each of these arguments is plausible, there is more reason to think that the conclusion is right.

Conduct, conclusions, and reasons will have different effects in different circumstances. We can imagine a group whose members are unimpressed by conclusions but much affected by behavior, or a group whose members pay far more attention to reasons than to conclusions.[106] By definition, the deliberative ideal is supposed to include reason-giving, not merely actions or statements of conclusions.[107] The problem is that when reasons are given, group members are likely to pay attention to them in a way that can lead such members to fail to say what they know. It follows that even when participants are acting rationally and in accordance with the deliberative ideal, they might well blunder. Indeed—and this is the central point—they blunder because of, rather than in spite of, their rationality and their fidelity to that ideal.

2. *Social Influences*

The second problem involves social influences. If people fear that their statements will be disliked or ridiculed, they might not speak out, even on questions of fact. Their silence might stem not from a belief that they are wrong, as in the case of informational pres-

---

104 For an overview exploring social pressures on individual judgments, see Solomon E. Asch, *Opinions and Social Pressure*, *in* READINGS ABOUT THE SOCIAL ANIMAL 13 (Elliott Aronson ed., 7th ed. 1995).

105 *See* DAVID KRECH ET AL., INDIVIDUAL IN SOCIETY 514 (1962) (showing individual susceptibility to majority views).

106 For relevant data, see Gene Rowe & George Wright, *Expert Opinions in Forecasting: The Role of the Delphi Technique*, *in* PRINCIPLES OF FORECASTING 125, 129–30 (J. Scott Armstrong ed., 2001).

107 *See* Habermas, *supra* note 31, at 940 (describing deliberative ideal).

Imaged with Permission from N.Y.U. Law Review

sure, but instead from the risk of social sanctions of various sorts. In the most extreme cases, those sanctions will take the form of criminal punishment (in societies that do not respect free speech) or complete exclusion from the group. In less severe cases, those who defy the dominant position will incur a form of disapproval that will lead them to be less trusted, liked, and respected in the future. Here, too, people are inevitably affected by the number and nature of those with the majority position. A large majority will impose more social pressure than a small one. If certain group members are leaders or authorities, willing and able to impose social sanctions of various sorts, others will be unlikely to defy them publicly. And if the group is especially cohesive, social pressures will be particularly intense. Hence the robust finding that "[h]ighly cohesive and group task-oriented groups tend to show a number of suboptimal decision-making symptoms, including intolerance of deviant ingroup opinions, censorship and self-censorship of deviants, and ultimately rejection of deviants."[108]

3. *A Framework:* Private Benefits vs. Social Benefits

Participation in deliberative processes, and the effects of informational and social influences, can be put into a more general framework. Suppose that group members are deliberating about some factual question; suppose too that each member has some information that bears on the answer to that question. Will members disclose what they know?

For each person, the answer may well depend on the individual benefits and the individual costs of disclosure.[109] In many situations, and entirely apart from informational and social influences, the individual benefits of disclosure will be far less than the social benefits. If I say what I know about a legal issue being examined by a team of lawyers, I will probably receive only a fraction of the benefit that comes from an improved decision by the group. And if each group member thinks this way, the group will receive only a fraction of the available information. In this sense, participants in deliberation face a standard collective action problem in which each person, following his rational self-interest, will tell the group less than it needs to know. At least this is so if each member receives only a small portion of the benefits that come to the group from a good outcome—a plausible view about the situation facing many institutions, including, for

---

[108] *See* José M. Marques et al., *Social Categorization, Social Identifiation, and Rejection of Deviant Group Members*, *in* GROUP PROCESSES, *supra* note 86, at 400, 403.

[109] I put altruistic motivations to one side; they bear on subsequent discussion, particularly the treatment of cascades in which members are rewarded for a good decision by the group, discussed *infra* Part II.E.

Imaged with Permission from N.Y.U. Law Review

example, corporate boards and administrative agencies. (Below I take up the question whether incentives might be restructured so as to remedy this problem, for instance by rewarding people for correct decisions by the group.[110])

If the statements of others suggest that privately held information is wrong or unhelpful, then the private benefit of disclosure is reduced even more. In that event, the group member has reason to believe that disclosure will not improve the group's decision at all. Things are even worse if those who speak against the apparent consensus will suffer reputational injury (or more). In that event, the private calculus is straightforward: Silence is golden. As we shall see, a great deal can be done to improve the situation by realigning individual incentives and through institutional design.

4. Findings

Both informational pressure and social influences help explain the finding that in a deliberating group, those with a minority position often silence themselves or otherwise have disproportionately little weight.[111] There is a more particular finding: Members of groups suffering from low social status—less educated people, sometimes women—speak less and carry less influence within deliberating groups than their higher-status peers.[112] Both informational influence and social pressures, likely to be especially strong for low-status members, contribute to this result. The unfortunate consequence can be a loss of information to the group as a whole, ensuring that deliberating groups do far less well than they would if only they could aggregate the information held by group members. If, in short, low-status members silence themselves, groups will fail, simply because such members often have relevant information.

Informational pressure and social pressures also help explain some otherwise puzzling findings about judicial voting on federal courts of appeals. Consider the fact that on three-judge panels, Republican appointees show far more conservative voting patterns when sitting with two other Republican appointees and that Democratic appointees show far more liberal voting patterns when sitting

110 *See infra* Part III.A.2.

111 *See* GLENN C. LOURY, SELF-CENSORSHIP IN PUBLIC DISCOURSE: A THEORY OF "POLITICAL CORRECTNESS" AND RELATED PHENOMENA 3 (Boston Univ., Ruth Pollak Working Paper Series on Economics, 1993).

112 *See* Caryn Christensen & Ann S. Abbott, *Team Medical Decision Making*, *in* DECISION MAKING IN HEALTH CARE 267, 272–76 (Gretchen B. Chapman & Frank A. Sonnenberg eds., 2000) (discussing effects of status on exchange of information in group interactions).

Imaged with Permission from N.Y.U. Law Review

with two other Democratic appointees.[113] Consider too the finding that, when sitting with two Republican appointees, Democratic appointees show quite conservative voting patterns, close to those of Republican appointees in the aggregate data; and that, when sitting with two Democratic appointees, Republican appointees are fairly liberal, with overall votes akin to those of Democratic appointees.[114] For federal judges, informational pressure and social influences are not the whole story, but they play a substantial role.[115]

More generally, a comprehensive study demonstrated that majority pressures can be powerful even for factual questions on which some people know the right answer.[116] The study involved 1200 people, forming groups of six, five, and four members.[117] Individuals were asked true-false questions, involving art, poetry, public opinion, geography, economics, and politics.[118] They were then asked to assemble into groups, which discussed the questions and produced answers. The views of the majority played a substantial role in determining the group's answers. If a majority of individuals in the group gave the right answer, the group's decision moved toward the majority in 79% of the cases.[119] The truth played a role too, but a lesser one. If a majority of individuals in the group gave the wrong answer, the group decision nonetheless moved toward the majority position in 56% of the cases.[120] Hence the truth did have an influence—79% is higher than 56%—but the majority's judgment was the dominant one. And because the majority was influential even when wrong, the average group decision was right only slightly more often than the average individual decision (66% vs. 62%).[121] What is most important is that groups did not perform as well as they would have if they had properly aggregated the information that group members possessed. The same basic outcome should be larger in highly cohesive groups with a shared sense of identity.[122]

---

113 *See* Sunstein, Schkade, & Ellman, *supra* note 23, at 314 (showing effects of panel composition on judicial decisions).

114 *See id.* at 305–11.

115 *See id.* at 337–46.

116 Robert L. Thorndike, *The Effect of Discussion Upon the Correctness of Group Decisions, When the Factor of Majority Influence Is Allowed For*, 9 J. SOC. PSYCHOL. 343, 348–61 (1938) (exploring effects of both correctness and majority pressure on group judgments).

117 *Id.* at 348.

118 *Id.* at 345.

119 *Id.* at 355.

120 *Id.*

121 *Id.* at 349.

122 *See* Joel Cooper et al., *Attitudes, Norms, and Social Groups*, *in* GROUP PROCESSES, *supra* note 86, at 259, 260–62.

Imaged with Permission from N.Y.U. Law Review

Most ambitiously, we might think that findings of this kind help to explain the rise of culture itself, and hence to illuminate cultural differences among groups and nations that might not be expected to be fundamentally different from one another.[123] If individuals affect one another, and if judgments and errors can spread from a few to many, then seemingly small differences are likely to be magnified through social influences. Even with respect to political issues, involving human rights and other issues, deliberation might produce significant cross-cultural differences as a result of variations in neighbors and starting points.[124] And if such differences emerge, there is no reason for confidence that good judgments are emerging from deliberation as such.

5. *Preconditions and the Internal Morality of Deliberation: A Hayekian Challenge to Habermas?*

It is now time to engage a broader question that I have raised throughout: Do these points amount to a challenge to deliberation as an ideal, or to deliberative conceptions of democracy? Many of those interested in deliberation have attempted to specify its preconditions in a way that is intended to ensure against some of the problems that I am emphasizing here. Jürgen Habermas, for example, stresses norms and practices designed to allow victory by "the better argument":

> Rational discourse is supposed to be public and inclusive, to grant equal communication rights for participants, to require sincerity and to diffuse any kind of force other than the forceless force of the better argument. This communicative structure is expected to create a deliberative space for the mobilization of the best available contributions for the most relevant topics.[125]

In Habermas's "ideal speech situation," all participants attempt to seek the truth; they do not behave strategically or attempt to decide; they accept a norm of equality.[126] Other advocates of deliberative democracy have spoken similarly about what appropriate deliberation entails.[127] On this view, deliberation, properly understood, does not simply involve the exchange of words and opinions. It

---

123 *See* Bibb Latané & Martin J. Bourgeois, *Dynamic Social Impact and the Consolidation, Clustering, Correlation, and Continuing Diversity of Culture*, *in* GROUP PROCESSES, *supra* note 86, at 235, 237–51.

124 *Id.* at 243–46.

125 *See* Habermas, *supra* note 31, at 940.

126 *See* JÜRGEN HABERMAS, *What is Universal Pragmatics?*, *in* COMMUNICATION AND THE EVOLUTION OF SOCIETY 1, 2–4, 32 (Thomas McCarthy trans., 1979) (discussing preconditions for communication).

127 *See* GUTMANN & THOMPSON, *supra* note 4, at 7–8 (outlining foundations of authors' vision of deliberative democracy).

Imaged with Permission from N.Y.U. Law Review

imposes its own requirements and preconditions. Indeed, deliberation has its own internal morality, one that operates as a corrective to some of the effects of deliberative processes in the real world.

These claims point in helpful directions. It is correct to say that deliberation, properly understood, contains an internal morality that can be invoked to challenge nominally deliberative processes. Habermas's preconditions will certainly make the situation better rather than worse, and better than it frequently is in real-world deliberative settings.

Unfortunately, preconditions of the sort identified by Habermas will cure few of the problems that I shall outline here. More particularly, those preconditions will do little to affect the four key failures on the part of deliberating groups. Each of the failures is likely to arise even if discourse is public and inclusive, even if participants are sincere, and even if everyone has equal communication rights. We might therefore take the argument to follow as a Hayekian critique of Habermas—a critique, that stresses (with Hayek) the diffusion of information in society and the difficulty of aggregating that information through deliberation (as opposed to the price signal).[128] As we shall see, some of the relevant problems are reduced if various forms of subtle "force" are eliminated. But the reduction is only partial. The four problems have distinctive structures; I discuss them in sequence.

C. Deliberative Failure 1: Amplification of Cognitive Errors

It is well known that individuals do not always process information well. They use heuristics that lead them to predictable errors; they are also subject to identifiable biases, which produce errors.[129] A growing literature explores the role of these heuristics and biases and their relationship to law and policy. For example, most people follow the representativeness heuristic, in accordance with which judgments of probability are influenced by assessments of resemblance (the extent to which A "looks like" B).[130] The representative heuristic helps explain what Paul Rozin and Carol Nemeroff call "sympathetic magical thinking," including the beliefs that some objects have conta-

---

[128] *See generally* F.A. Hayek, *The Use of Knowledge in Society*, 35 AM. ECON. REV. 519 (1945) (discussing dispersal of knowledge and its aggregation through markets).

[129] For an overview, see generally HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT (Thomas Gilovich et al. eds., 2002). *See also* BEHAVIORAL LAW & ECONOMICS (Cass R. Sunstein ed., 2000).

[130] Amos Tversky & Daniel Kahneman, *Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment*, *in* HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 129, at 19, 22–25 (discussing representative conjunctions).

Imaged with Permission from N.Y.U. Law Review

gious properties and that causes resemble their effects.[131] The representativeness heuristic often works well, but it can also lead to severe blunders.

People also err because they use the availability heuristic to answer difficult questions about probability. When people use this heuristic, they answer a question of probability by asking whether examples come readily to mind.[132] In addition, most people are strikingly vulnerable to framing effects, making different decisions depending on the wording of the problem. For a simple example, consider the question whether to undergo a risky medical procedure. When people are told, "Of those who have this procedure, 90% are alive after five years," they are far more likely to agree to the procedure than when they are told, "Of those who have this procedure, 10% are dead after five years."[133]

For purposes of assessing deliberation, a central question is whether groups avoid the errors of the individuals who compose them. There is no clear evidence that they do, and there is considerable evidence that they do not—a vivid illustration of the principle, "garbage in, garbage out," in a way that mocks the aspiration to collective correction of individual blunders. In fact individual errors are not merely replicated but actually amplified in group decisions—a process of "some garbage in, much garbage out." The most general finding is that when a bias is widely shared, group interactions will actually enhance its effect.[134]

Consider some particular findings. If individual jurors are biased because of pretrial publicity that misleadingly implicates the defendant, or even because of the defendant's unappealing physical appearance, juries as a group are likely to amplify rather than to correct those biases.[135] Groups have been found to amplify, rather than to attenuate, reliance on the representativeness heuristic;[136] to fall prey

131 Paul Rozin & Carol Nemeroff, *Sympathetic Magical Thinking: The Contagion and Similarity "Heuristics," in* HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT, *supra* note 129, at 201–16 (exploring perceptions involving contagion).

132 *See* Amos Tversky & Daniel Kahneman, *Availability: A Heuristic For Judging Frequency and Probability*, 5 COGNITIVE PSYCHOL. 207, 208 (1973) (discussing availability heuristic).

133 *See* Donald A. Redelmeier et al., *Understanding Patients' Decisions: Cognitive and Emotional Perspectives*, 270 JAMA 72, 73 (1993) (discussing framing effects in medical context).

134 *See* Stasser & Dietz-Uhler, *supra* note 86, at 49–50. Note that when the bias is not widely shared, it may be corrected through deliberation. *See id.*

135 MacCoun, *supra* note 10, at 121–26 (showing amplification of jury bias).

136 Mark F. Stasson et al., *Group Consensus Approaches on Cognitive Bias Tasks: A Social Decision Scheme Approach*, 30 JAPANESE PSYCHOL. RES. 68, 74–75 (1988).

Imaged with Permission from N.Y.U. Law Review

to even larger framing effects than individuals;[137] to show more overconfidence than group members;[138] to be more affected by the biasing effect of spurious arguments from lawyers;[139] to be more susceptible to the "sunk cost fallacy";[140] and to be more subject to choice-rank preference reversals.[141] In an especially revealing finding, groups have been found to make more, rather than fewer, conjunction errors (believing that A and B are more likely to be true than A alone) than individuals when individual error rates are high—though fewer when individual error rates are low.[142] In addition, groups demonstrate only a marginally decreased level of reliance on the availability heuristic, even when use of that heuristic leads to clear errors.[143] In a particularly disturbing finding, it is shown that groups are more likely than individuals to escalate their commitment to a course of action that is failing, and all the more so if members identify strongly with the groups of which they are a part.[144]

Why are individual cognitive errors propagated and often amplified at the group level? Informational pressures and social influences are unquestionably at work. Suppose, for example, that most members of a group are prone to make conjunction errors.[145] If the majority makes conjunction errors, then most people will see others making conjunction errors, and what they see will convey information about what is right. Those who are not specialists in logic are likely to think: If most people make conjunction errors, perhaps they are not errors at all. Of course some people will not fall prey to those errors

---

137 *See* Kerr et al., *supra* note 64, at 693, 711–12.

138 *See id.* at 692 tbl.1 (noting study that found groups generally more confident than individuals); Janet A. Sniezek & Rebecca A. Henry, *Accuracy and Confidence in Group Judgment*, 42 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 1, 24–27 (1989). This finding very much bears on excessive risk-taking, including in the context of making war. *See generally* JOHNSON, *supra* note 25.

139 *See* Kerr et al., *supra* note 64, at 691.

140 The sunk cost fallacy ("throwing good money after bad") emerges when people reason, after making a bad investment: I shouldn't stop now, because if I do, I will lose what I have already paid out. On the sunk cost fallacy within groups, see generally Glen Whyte, *Escalating Commitment in Individual and Group Decision Making: A Prospect Theory Approach*, 54 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 430 (1993).

141 *See* John C. Mowen & James W. Gentry, *Investigation of the Preference-Reversal Phenomenon in a New Product Introduction Task*, 65 J. APPLIED PSYCHOL. 715, 721 (1980). *But see* Julie R. Irwin & James H. Davis, *Choice/Matching Preference Reversals in Groups: Consensus Processes and Justification Based Reasoning*, 64 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 325, 337 (1995) (finding preference reversals to have been moderated by group discussion).

142 Kerr et al., *supra* note 64, at 692 (citing studies).

143 *Id.*

144 Stasser & Dietz-Uhler, *supra* note 86, at 48.

145 *See* Tversky & Kahneman, *supra* note 130, at 26.

Imaged with Permission from N.Y.U. Law Review

and may even correct them; but group members would have to have a high degree of confidence to do so. Recall here the finding that groups make more conjunction errors than individuals when the initial rate of individual error is high[146]—a finding that fits well with the informational explanation of why groups amplify errors.

Social influences also contribute to the propagation and amplification of individual mistakes. If most group members make conjunction errors, others also might make them simply in order not to seem disagreeable or foolish—at least if there is no particular incentive to produce the right answer. And if most group members use the availability heuristic, or commit the sunk-cost fallacy, then there will be social pressure for the other members to do the same.

To be sure, there is some evidence of group attenuation of certain biases. For example, groups are slightly less susceptible to hindsight bias.[147] Apparently members who are not susceptible to that bias are able to persuade others that it is indeed a bias. Groups are especially likely to outperform the average individual when members are subject to "egocentric biases."[148] When asked what percentage of other undergraduates will vote for a particular candidate, have cell phones, watch television on Tuesday night, enjoy a particular singer, or believe that the latest Spiderman movie will win at least one Oscar, most people show a bias in the direction that they themselves favor. They believe that their tastes and preferences are typical. But in groups with diverse views, individual members learn that their own position is not universally held, and hence the bias is reduced.[149] In these cases, group deliberation supplies an important corrective.

But the broader point is that, with group discussion, individual errors are usually propagated, not eliminated,[150] and amplification of mistakes is as likely as alleviation. A general review suggests that when individuals show a high degree of bias, groups are likely to be more biased, not less biased, than their median or average member; in such circumstances, "groups generally can be expected to amplify rather than correct individual bias."[151] This point is an application of the lesson, from the Condorcet Jury Theorem, that as the size of the group expands, the likelihood of group error expands toward 100% if

146 *See* Kerr et al., *supra* note 64, at 692.

147 *See generally* Dagmar Stahlberg et al., *We Knew It All Along: Hindsight Bias in Groups*, 63 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 46 (1995).

148 Personal communication with Reid Hastie, University of Chicago Business School (July 24, 2004), who has conducted experiments on this issue for many years.

149 *Id.*

150 *See* Bottom et al., *supra* note 52, at 160.

151 MacCoun, *supra* note 10, at 124 (emphasis omitted).

Imaged with Permission from N.Y.U. Law Review

each group member is more likely to be wrong than right. What I am emphasizing here is that social dynamics can aggravate rather than reduce that problem. And if this is so, then jury deliberations, as well as deliberation within multimember courts and the executive branch, will be prone to error. Recall here the suggestion that both the CIA and NASA blundered because group processes failed to correct, and instead amplified, initial biases internal to both agencies.[152]

D. Deliberative Failure 2: Hidden Profiles and Common Knowledge

Suppose that group members have a great deal of information—enough to produce the unambiguously correct outcome if that information is properly aggregated. Even if this is so, an obvious problem is that groups will not perform well if they emphasize shared information and slight information that is held by only one or a few members. Unfortunately, countless studies demonstrate that this unfortunate result is highly likely.[153] "Hidden profiles" is the term for accurate understandings that groups could obtain but do not. Hidden profiles are in turn a product of the "common knowledge effect," through which information held by all group members has more influence on group judgments than information held by only a few members.[154] The most obvious explanation of the effect is the simple fact that, as a statistical matter, common knowledge is more likely to be communicated to the group; but social influences play a role as well.

1. Examples

Consider a study of serious errors within working groups, both face-to-face and online.[155] The purpose of the study was to see how groups might collaborate to make personnel decisions. Resumes for three candidates applying for a marketing manager position were placed before group members.[156] The attributes of the candidates were manipulated by the experimenters so that one applicant was clearly the best candidate for the job described. Packets of information were given to subjects, each containing a subset of information

---

152 *See supra* notes 12–19 and accompanying text.

153 *See* Stasser & Titus, *Hidden Profiles*, *supra* note 10, at 304, 306–13 (2003) (discussing hidden profile experiments).

154 Daniel Gigone & Reid Hastie, *The Common Knowledge Effect: Information Sharing and Group Judgments*, 65 J. PERSONALITY & SOC. PSYCHOL. 959, 971–73 (1993) [hereinafter Gigone & Hastie, *Common Knowledge Effect*] (explaining hidden profiles by reference to common knowledge effect).

155 *See* Ross Hightower & Lutfus Sayeed, *The Impact of Computer-Mediated Communication Systems on Biased Group Discussion*, 11 COMPUTERS HUM. BEHAV. 33, 43 (1995).

156 *Id.* at 39.

Imaged with Permission from N.Y.U. Law Review

from the resumes, so that each group member had only part of the relevant information. The groups consisted of three people, some operating face-to-face, some operating online. Almost none of the deliberating groups made what was conspicuously the right choice.[157] The reason is simple: Members failed to share information in a way that would permit the group to make that choice. People were inclined to share positive information about the winning candidate and negative information about the losers. They suppressed negative information about the winner and positive information about the losers. Hence their statements served to "reinforce the march toward group consensus rather than add complications and fuel debate."[158]

Or consider a simulation of political elections, in which information was parceled out to individual members about three candidates for political office, and in which properly pooled information could have led to the selection of Candidate A, who was clearly the best choice.[159] In the first condition, each member of the four-person group was given most of the relevant information (66% of the information about each candidate). In that condition, 67% of group members favored Candidate A before discussion, and 85% after discussion.[160] This is a clear example of appropriate aggregation of information. Groups significantly outperformed individuals, apparently because of the exchange of information and reasons. Here, then, is a clear illustration of the possibility that groups can aggregate what members know in a way that produces sensible outcomes.

In the second condition, by contrast, the information that favored Candidate A was parceled out to various members of the group, so that only 33% of information about each candidate was shared, and 67% was unshared. As the condition was designed, the shared information favored two unambiguously inferior candidates, B and C.[161] If the unshared information emerged through discussion, and was taken seriously, Candidate A would be chosen. In that condition, less than 25% of group members favored Candidate A before discussion, a natural product of the initial distribution of information. But (and this is the key result) the number favoring Candidate A actually *fell* after discussion, simply because the shared information had dispropor-

---

157 *Id.* at 40.

158 PATRICIA WALLACE, THE PSYCHOLOGY OF THE INTERNET 82 (1999).

159 *See* Garold Stasser & William Titus, *Pooling of Unshared Information in Group Decision Making: Biased Information Sampling During Discussion*, 48 J. PERSONALITY & SOC. PSYCHOL. 1467, 1471–72 (1985) [hereinafter Stasser & Titus, *Pooling*].

160 *Id.* at 1473; *see also* Stasser & Titus, *Hidden Profiles*, *supra* note 10, at 304.

161 Stasser & Titus, *Pooling*, *supra* note 159, at 1471–72.

Imaged with Permission from N.Y.U. Law Review

tionate influence on group members.[162] In other words, groups did worse, not better, than individuals when the key information was distributed selectively. In those conditions, the commonly held information was far more influential than the unshared information, to the detriment of the group's ultimate decision.

From this and many similar studies, the general conclusion is that when "the balance of unshared information opposes the initially most popular position. . . . the unshared information will tend to be omitted from discussion and, therefore, will have little effect on members' preferences during group discussion."[163] That conclusion has a clear connection with the judgments, mentioned above, about large-scale information failures at the CIA and similar failures at NASA.[164] It follows that "[g]roup decisions and postgroup preferences reflect[ ] the initial preferences of group members even when the exchange of unshared information should have resulted in substantial shifts in opinion."[165] Nor does discussion increase the recall of unshared information. Instead, its major effect is to increase recall of the attributes of the initially most popular candidate.[166] The most disturbing conclusion is that when key information is unshared, groups are "more likely to endorse an inferior option after discussion than . . . their individual members before discussion."[167]

2. *The Common Knowledge Effect*

These results are best understood as a consequence of the "common knowledge effect," by which information held by all group members has the most substantial influence on group judgments, far more than information held by one member or a few.[168] More precisely, the "influence of a particular item of information is directly and positively related to the number of group members who have knowledge of that item before the group discussion and judgment."[169] Under conditions of unshared information, group judgments have been found to be "not any more accurate than the average of the individual judgments, even though"—and this is the central point—"the

---

162 *Id.* at 1473 tbl.3.

163 *Id.* at 1476.

164 *See supra* notes 12–19 and accompanying text.

165 Stasser & Titus, *Pooling*, *supra* note 159, at 1476.

166 *Id.*

167 Stasser & Titus, *Hidden Profiles*, *supra* note 10, at 305.

168 *See* Gigone & Hastie, *Common Knowledge Effect*, *supra* note 154, at 959 (describing experiment showing common knowledge effect in groups of three).

169 *Id.* at 960.

Imaged with Permission from N.Y.U. Law Review

groups were in possession of more information than were any of the individuals."[170]

In a key study, deliberating groups would have lost nothing in terms of accuracy if they had simply averaged the judgments of the people involved—a clear finding that deliberation may not improve on the judgments of statistical groups.[171] The more shared information is (that is, the more that it stands as "common knowledge"), the more impact it will have on group members before discussion begins—and the more impact it will have as discussion proceeds, precisely because commonly held information is more likely to be discussed.

As might be expected, the group's focus on shared information increases with the size of the group.[172] In another study designed to test judgments about candidates for office, involving both three-person and six-person groups, all discussions focused far more on shared information than on unshared information—but the effect was significantly greater for six-person groups. Most remarkably, "it was almost as likely for a shared item to be mentioned twice as it was for an unshared item to be mentioned at all."[173] And despite the failures of their deliberations, group members were significantly more confident in their judgments after discussion.[174]

How can these findings be squared with the Condorcet Jury Theorem? The most fundamental point is that in deliberation, individuals are not making judgments on their own; they are being influenced by the judgments of others. When interdependent judgments are being made, and when some people are wrong, the Condorcet Jury Theorem offers no clear predictions. Under such circumstances, it is not at all clear that groups will reach better conclusions than individuals.[175] And when groups fail, the tendency toward hidden profiles is often part of the reason.

---

170 *Id.* at 973.

171 *Id.*

172 *See* Garold Stasser et al., *Information Sampling in Structured and Unstructured Discussions of Three- and Six-Person Groups*, 57 J. PERSONALITY & SOC. PSYCHOL. 67, 72–73 (1989).

173 *Id.* at 78.

174 *Id.* at 72.

175 *See generally* Kerr et al., *supra* note 64, at 714 (predicting enhanced bias by groups facing "many real-world decision tasks" and listing as examples, "jury decision making, hiring decisions, risky investment decisions" and certain foreign policy decisions). On some of the theoretical issues, see David Austen-Smith & Jeffrey S. Banks, *Information Aggregation, Rationality, and the Condorcet Jury Theorem*, 90 AM. POL. SCI. REV. 34 (1996).

Imaged with Permission from N.Y.U. Law Review

3. *Informational Influences and Social Pressures Redux*

Why do hidden profiles remain hidden? The two major explanations track the informational and social accounts traced above. When information is held by all or most group members, it is especially likely, as a statistical matter, to be repeated in group discussion, and hence more likely to be influential than information that is held by one person or a few.[176] There are two different points here. First, information held by all or most group members is likely to influence individual judgments, and those judgments will affect the judgments of the group.[177] Thus the effects of a shared piece of information will influence the group simply through its impact on predeliberation judgments. Second, shared information, because it is shared, is more likely to be explored during group discussion.[178]

Suppose, for example, that a team of five lawyers is deciding whether to appeal an adverse trial court ruling. If each of the five lawyers has information indicating that an appeal would be unsuccessful, that information is more likely to emerge in group discussion than separate parcels of information, individually held by each lawyer, suggesting that an appeal would succeed. If the team of lawyers stresses the information that is antecedently held by each, that information will have a disproportionate influence on its ultimate decision.[179] This is a statistical point about information sampling.

But information sampling provides an incomplete account; hidden profiles remain even more hidden than would be predicted by that account.[180] To understand the additional element, consider the finding that low-status members of groups are "increasingly reluctant over the course of discussion to repeat unique information."[181] Those in a group who are inexperienced, or are thought to be low on the hierarchy, are particularly loathe to emphasize their privately held information as discussion proceeds. This finding suggests that group members, and especially lower status ones, are alert to the reputational costs of emphasizing information that most group members seem to lack. Lower status members "are likely to drop unique information like a hot potato"—partly because of the difficulty of establishing its credibility and relevance, and partly because they may incur group disapproval if they press a line of argument that others reject.[182]

---

176 *See* Stasser & Titus, *Hidden Profiles*, *supra* note 10, at 306–07.
177 *See* Gigone & Hastie, *Common Knowledge Effect*, *supra* note 154, at 960.
178 *Id.*
179 *Id.*
180 Stasser & Titus, *Hidden Profiles*, *supra* note 10, at 308.
181 *Id.*
182 *Id.*

Imaged with Permission from N.Y.U. Law Review

With respect to the risk of error, consider the finding that group members underestimate the performance of low-status members and overestimate the performance of high-status members, in a way that gives high-status members a degree of deference that is not warranted by reality.[183] It follows that hidden profiles are produced by both informational and reputational pressures imposed by the initial distribution of views.

In the same vein, those who discuss shared information obtain rewards in the form of an enhanced sense of competence and standing in the eyes of others—and in their own eyes as well.[184] In both face-to-face discussions and purely written tasks, people give higher ratings (in terms of knowledge, competence, and credibility) to themselves and to others after receiving information that they knew already. It follows that "a bearer of valuable, unshared information may need to establish credibility by telling others what they already know before telling them what they do not already know."[185] The general problem is that deliberating groups often perform poorly because they fail to elicit information that could steer them in the right directions.

E. Deliberative Failure 3: Cascades

1. Informational Cascades

Hidden profiles are closely related to informational cascades, which greatly impair group judgments. Cascades need not involve deliberation, but deliberative processes often involve cascades. As in the case of hidden profiles, the central point is that those involved in a cascade do not reveal what they know. As a result, the group does not obtain important information.

To see how informational cascades work, imagine a deliberating jury that is deciding whether a defendant should be subject to a punitive damage award and, if so, in what amount.[186] Let us also assume that the jurors are announcing their views in sequence, in a temporal queue, and that each juror knows his place on that queue. From his own recollection of the evidence and the jury instructions, and from

---

183 Cecilia L. Ridgeway, *Social Status and Group Structure*, *in* GROUP PROCESSES, *supra* note 86, at 352, 354 (collecting studies).

184 *See* Gwen M. Wittenbaum et al., *Mutual Enhancement: Toward an Understanding of the Collective Preference for Shared Information*, 77 J. PERSONALITY & SOC. PSYCHOL. 967, 967–78 (1999).

185 Stasser & Titus, *Hidden Profiles*, *supra* note 10, at 311.

186 I draw here on David Hirshleifer, *The Blind Leading the Blind: Social Influence, Fads, and Informational Cascades*, *in* THE NEW ECONOMICS OF HUMAN BEHAVIOR 188, 193–95 (Mariano Tommasi & Kathryn Ierulli eds., 1995), and on the discussion in SUNSTEIN, *supra* note 21, at 55–73 (2003).

Imaged with Permission from N.Y.U. Law Review

some personal experience, each juror has some private information about what should be done. But each juror also attends, reasonably enough, to the judgments of others. Andrews is the first to speak. He suggests that the defendant should be subject to a punitive award and a high one—say, $5 million. Barnes now knows Andrews's judgment; it is clear that she too should certainly urge a punitive award, and a high one, if she agrees independently with Andrews. But if her independent judgment is that no award should be imposed, she would—if she trusts Andrews no more and no less than she trusts herself—be indifferent about what to do, and might simply flip a coin.

Now turn to a third juror, Carlton. Suppose that both Andrews and Barnes have favored a punitive award, and a multimillion dollar one, but that Carlton's own information, though inconclusive, suggests that no award should be imposed. In that event, Carlton might well ignore what he knows and follow Andrews and Barnes. It is likely, in these circumstances, that both Andrews and Barnes had reasons for their conclusion, and unless Carlton thinks that his own information is better than theirs, he should follow their lead. If he does, Carlton is in a cascade. Now suppose that Carlton is acting in response to what Andrews and Barnes did, not on the basis of his own information, and also that subsequent jurors know what Andrews, Barnes, and Carlton did. On reasonable assumptions, they will do exactly what Carlton did: favor a high punitive damage award regardless of their private information (which, we are supposing, is relevant but inconclusive). This will happen even if Andrews initially blundered.[187]

If this is what is happening, there is a serious social problem: Jurors who are in the cascade do not disclose the information that they privately hold. In the example just given, jury decisions will not reflect the overall knowledge, or the aggregate knowledge, of those on the jury—even if the information held by individual jurors, if actually revealed and aggregated, would produce a quite different result. The reason for the problem is that individual jurors are following the lead of those who came before. Subsequent jurors might fail to rely on, and fail to reveal, private information that actually exceeds the information collectively held by those who started the cascade.

Cascades often occur in the real world within deliberating groups and in other situations;[188] they are easy to create in the laboratory as well. The simplest experiment asked subjects to guess whether the experiment used Urn A, which contained two white balls and one

187 *See generally* Hirshleifer, *supra* note 186, at 193–95.

188 *See id.* at 200–07 (discussing real world examples of cascades); *see also* SUNSTEIN, *supra* note 21, at 54–95 (discussing cascades).

Imaged with Permission from N.Y.U. Law Review

dark, or Urn B, which contained two dark balls and one white.[189] Subjects could earn $2 for a correct decision, and hence an economic incentive favored correct individual decisions (a point to which I will return). In each round, the contents of the chosen urn were emptied into a container. A randomly selected subject was asked to make one (and only one) private draw of a ball. After that draw, the subject recorded, on an answer sheet, the color of the draw and her own decision about which urn was involved. The subject did not announce her draw to the group, but she did announce her own *decision* to everyone. Then the urn was passed to the next subject for her own private draw, which again was not disclosed, and for her own decision about the urn, which again was disclosed.[190] This process continued until all subjects had made draws and decisions. At that time the experimenter announced the actual urn used. If the subject picks the urn based only on her private information, she will be right 66.7% of the time. The point of the experiment is to see whether people will decide to ignore their own draw in the face of conflicting announcements by predecessors—and to explore whether such decisions will lead to cascades and errors.

In the experiment, cascades often developed, and they often produced errors. After a number of individual judgments were revealed, people sometimes announced decisions that were inconsistent with their private draws, but that fit with the majority of previous announcements. Over 71% of "rounds" resulted in cascades.[191] Consider cases in which one person's draw (say, white ball) contradicted the announcement of his predecessor (say, Urn B). A study that replicated the urn experiment found that the second announcement nonetheless matched the first about 11% of the time—far less than a majority, but enough to ensure cascades.[192] And when one person's draw contradicted the announcement of two or more predecessors, the second announcement was likely to follow those who went before. Notably, the majority of decisions were rationally based on the available information[193]—but erroneous cascades nonetheless developed.

---

189 *See* Lisa R. Anderson & Charles A. Holt, *Information Cascades in the Laboratory*, 87 AM. ECON. REV. 847, 849–53, 860 (1997) (exploring experimental evidence of cascades).

190 *Id.* at 851.

191 *Id.* at 859.

192 Angela A. Hung & Charles R. Plott, *Information Cascades: Replication and an Extension to Majority Rule and Conformity-Rewarding Institutions*, 91 AM. ECON. REV. 1508, 1518 (2001).

193 *See* Anderson & Holt, *supra* note 189, at 853. A majority of subjects also rationally followed available information in Marc Willinger & Anthony Ziegelmeyer, *Are More Informed Agents Able to Shatter Information Cascades in the Lab*, *in* THE ECONOMICS OF NETWORKS: INTERACTION & BEHAVIOURS 291, 304 (Patrick Cohendet et al. eds., 1998).

Imaged with Permission from N.Y.U. Law Review

Here is an actual example of a cascade producing an inaccurate outcome (the urn used was B):[194]

TABLE 1: AN INFORMATIONAL CASCADE

	1	2	3	4	5	6
Private Draw	A	A	B	B	B	B
Decision	A	A	A	A	A	A

What is noteworthy here, of course, is that the total amount of private information—four darks and two whites—justified the correct judgment, which was in favor of Urn B. But the existence of two early signals, producing rational but incorrect judgments, led everyone else to fall in line. "[I]nitial misrepresentative signals start a chain of incorrect decisions that is not broken by more representative signals received later."[195] This result maps directly onto real-world decisions by deliberating groups, in which people fail to disclose what they know, to the detriment of the group as a whole. As a possible example, consider the existence of widely divergent group judgments about the origins and causes of AIDS, with some groups believing, falsely, that the first cases were observed in Africa as a result of sexual relations with monkeys, and with other groups believing, also falsely, that the virus was produced in government laboratories.[196] These and other views about AIDS are a product of social interactions and in particular of cascade effects.

2. *Reputational Cascades*

In a reputational cascade, people think that they know what is right, or what is likely to be right, but they nonetheless go along with the crowd in order to maintain the good opinion of others. Suppose that Albert suggests that global warming is a serious problem, and that Barbara concurs with Albert, not because she actually thinks that Albert is right, but because she does not wish to seem, to Albert, to be ignorant or indifferent to environmental protection. If Albert and Barbara seem to agree that global warming is a serious problem, Cynthia might not contradict them publicly and might even appear to share their judgment, not because she believes that judgment to be correct, but because she does not want to face their hostility or lose their good opinion.

---

194 *See* Willinger & Ziegelmeyer, *supra* note 193, at 291.

195 Anderson & Holt, *supra* note 189, at 859.

196 *See* Fabio Lorenzi-Cioldi & Alain Clémence, *Group Processes and the Construction of Social Representations*, *in* GROUP PROCESSES, *supra* note 86, at 311, 315–17.

Imaged with Permission from N.Y.U. Law Review

It should be easy to see how this process might generate a cascade. Once Albert, Barbara, and Cynthia offer a united front on the issue, their friend David might be most reluctant to contradict them even if he thinks that they are wrong. The apparent views of Albert, Barbara, and Cynthia carry information; that apparent view might be right. But even if David thinks that they are wrong and has information supporting that conclusion, he might be most reluctant to take them on publicly. In the actual world of group decisions, people are of course uncertain whether publicly expressed statements are a product of independent knowledge, participation in an informational cascade, or reputational pressure. Much of the time, listeners and observers probably overstate the extent to which the actions of others are based on independent information.

The possibility of reputational cascades is demonstrated by an ingenious variation on the urn experiment described above.[197] In this experiment, people were paid twenty-five cents for a correct decision, but seventy-five cents for a decision that matched the decision of the majority of the group. There were punishments for incorrect and nonconforming answers as well. If people made an incorrect decision, they lost twenty-five cents; if their decision failed to match the group's decision, they lost seventy-five cents.[198]

In this experiment, cascades appeared almost all of the time. No fewer than 96.7% of rounds resulted in cascades, and 35.3% of people's announcements did not match their private signal, that is, the signal given by their own draw.[199] And when the draw of a subsequent person contradicted the announcement of the predecessor, 72.2% of people matched the first announcement.[200] Consider, as a dramatic illustration, this round of the experiment[201] (the actual urn for this round was B):

TABLE 2: CONFORMITY AND CASCADES

	1	2	3	4	5	6	7	8	9	10
Private Draw	A	B	B	B	A	B	B	B	A	B
Decision	A	A	A	A	A	A	A	A	A	A

197 *See* Hung & Plott, *supra* note 192, at 1511–12 (offering variations on urn experiment).

198 *Id.*

199 *Id.* at 1517–18.

200 *Id.* at 1518.

201 *Id.* at 1516 (conformity-rewarding model, experiment 4, period 10).

Imaged with Permission from N.Y.U. Law Review

This experiment shows that especially unfortunate results should be expected if people are rewarded not only or not mostly for being correct, but also or mostly for doing what other people do.

F. Deliberative Failure 4: Group Polarization

There are clear links among hidden profiles, social cascades, and the well-established phenomenon of group polarization, by which *members of a deliberating group end up in a more extreme position in line with their tendencies before deliberation began.*[202] Group polarization is the typical pattern with deliberating groups. It has been found in hundreds of studies involving over a dozen countries.[203] For example, those who disapprove of the United States, and are suspicious of its intentions, will increase their disapproval and suspicion if they exchange points of view. Indeed, there is specific evidence of this phenomenon among citizens of France.[204]

Group polarization occurs for issues of fact as well as issues of value, though it is easier to demonstrate for the latter.[205] Group polarization has been found on obscure factual questions, such as how far Sodom (on the Dead Sea) is below sea level.[206] But if the question is whether a terrorist attack will occur in the United States in the next year, group polarization will not be easy to test, simply because the answer is either yes or no, and it is not simple to demonstrate greater extremism in binary choices. But suppose that people are asked, on a bounded scale of zero to eight, how likely it is that a terrorist attack will occur in the United States in the next year, with zero indicating "zero probability," eight indicating "absolutely certain," seven indicating, "overwhelmingly likely," six "more probable than not," and five "fifty-fifty." In that event, the answers from a deliberating group will tend to reveal group polarization, as people move toward more extreme points on the scale, depending on their initial median point. If the predeliberation median is six, the group judgment will usually be seven; if the predeliberation median is three, the group judgment will usually be two.[207] Recall here that federal judges are highly susceptible to group polarization, as both Democratic and Republican appointees show far more ideological voting patterns when sitting

---

202 *See* BROWN, *supra* note 22, at 202–26.

203 *See id.* at 204.

204 *Id.* at 223–24.

205 *See* JOHN C. TURNER, REDISCOVERING THE SOCIAL GROUP: A SELF-CATEGORIZATION THEORY 152–53 (1987).

206 *Id.*

207 *See* BROWN, *supra* note 22, at 222–24.

Imaged with Permission from N.Y.U. Law Review

with other judges appointed by a president of the same political party.[208] Juries polarize as well.[209]

Why does group polarization occur? There are three reasons.[210] The first and most important involves the now-familiar idea of informational influence, but in a distinctive form. People respond to the arguments made by other people—and the "argument pool," in any group with some predisposition in one direction, will inevitably be skewed toward that predisposition. As a statistical matter, the arguments favoring the initial position will be more numerous than the arguments pointing in the other direction. Individuals will have heard of some, but not all, of the arguments that emerge from group deliberation. As a result of the relevant arguments, deliberation will lead people toward a more extreme point in line with what group members initially believed.

The second explanation involves social influences.[211] People want to be perceived favorably by other group members. Sometimes people's publicly stated views are, to a greater or lesser extent, a function of how they want to present themselves. Once they hear what others believe, some will adjust their positions at least slightly in the direction of the dominant position, to present themselves in the way that they prefer.[212]

The third explanation stresses that people with extreme views tend to have more confidence that they are right, and that as people gain confidence, they become more extreme in their beliefs.[213] In a wide variety of experimental contexts, people's opinions have been shown to become more extreme simply because their views have been corroborated, and because they have become more confident after learning of the shared views of others.[214]

A great deal of work suggests that group polarization is heightened when members have a sense of shared identity, and this point is sometimes used to suggest an independent explanation of polariza-

---

208 *See* Sunstein, Schkade & Ellman, *supra* note 23, at 305 (showing group polarization within court of appeals panels).

209 *See* David Schkade et al., *Deliberating About Dollars: The Severity Shift*, 100 COLUM. L. REV. 1139, 1140–41 (2000) (showing group polarization with mock juries).

210 *See* BROWN, *supra* note 22, at 212–22, 226–45; Baron et al., *supra* note 83, at 540.

211 BROWN, *supra* note 22, at 233–39.

212 *Id.* at 213–17. It similarly has been suggested that majorities are especially potent because people do not want to incur the wrath, or lose the favor, of large numbers of people, and that when minorities have influence, it is because they produce genuine attitudinal change. *See* Baron et al., *supra* note 83, at 550, 557–59.

213 *See* Baron et al., *supra* note 83, at 557–59 (showing that corroboration increases confidence and hence extremism).

214 *Id.* at 541, 546–47, 557 (concluding that corroboration of one's views has effects on opinion extremity).

Imaged with Permission from N.Y.U. Law Review

tion, involving intergroup differentiation.[215] People may polarize because they are attempting to conform to the position that they see as typical within their own group; if in-group identity is especially salient or important, the in-group norms "are likely to become more extreme so as to be more clearly differentiated from outgroup norms, and the within-group polarization will be enhanced."[216] When Democrats or Republicans become polarized, or when polarization occurs within religious or ethnic groups, intergroup differentiation is likely to be a major reason. And if arguments come from a member of an in-group, they are especially likely to be persuasive, and it is reasonable to think that people would fear the reputational sanctions that come from rejecting what an in-group member has to say. By contrast, the views of out-group members have less force.[217] The clear lesson is that when a group is highly cohesive, and when members are closely identified with it, polarization is especially likely and likely to be especially large.

Does group polarization lead to accurate or inaccurate answers? Do deliberating groups err when they polarize? No general answer would make sense. Everything depends on the relationship between the correct answer and the group's predeliberation tendencies. But as a result of the relevant influences, some people will fail to disclose what they know. Deliberative processes might well fail to move people in the right directions. When individuals are leaning in a direction that is mistaken, the mistake will be amplified by group deliberation. We have already encountered an example: When most people are prone to make conjunction errors, group processes lead to more errors rather than fewer.[218] The same is true when jury members are biased as a result of pretrial publicity; here the jury as a group becomes more biased than individual jurors were.[219] This is polarization in action, and it produces major blunders.

G. Deliberative Success?

Thus far I have emphasized several reasons why deliberation often fails to improve on the judgments of statistical groups, and indeed might make those judgments even worse. But there is some intriguing countervailing evidence.

---

[215] *See* BROWN, *supra* note 82, at 209–11; TURNER, *supra* note 205, at 159–70; Joel Cooper et al., *supra* note 122, at 259, 269–70.

[216] BROWN, *supra* note 82, at 210.

[217] *Id.* at 211; Cooper et al., *supra* note 122, at 269.

[218] *See supra* notes 134–44 and accompanying text.

[219] *See* MacCoun, *supra* note 10, at 127–28.

Imaged with Permission from N.Y.U. Law Review

1. *Increases in Accuracy*

When one or more people in a group are confident that they know the right answer to a factual question, the group might be expected to shift in the direction of accuracy.[220] And if the question has a readily demonstrable answer, it is more likely that groups will converge on it.[221] For problems with answers that are self-affirming, or that are clearly right once stated, a single correct member usually ensures a correct answer from the group.[222] Suppose that the question is how many people were on the earth in 1940, or the number of Supreme Court decisions invalidating acts of Congress, or the distance between Paris and London. Suppose too that one or a few people know the right answer. If so, there is a good chance that the group will not polarize, but instead accept that answer.[223]

When this is so, the reason is simple: The person who is confident that she knows the answer will speak with assurance and authority, and she is likely to be convincing for that very reason. An early study finds that those with correct answers are usually more confident, and hence confidence was "associated with correctness for both individual and group performance."[224] Consider in this light the finding that pairs tend to do better than individuals on a test involving general vocabulary knowledge; those pairs with at least one high-ability member generally performed at the same level as their more competent member.[225]

Some evidence suggests that while deliberating groups often fail to spread information, they are less likely to neglect unshared information if they believe that there is a demonstrably correct answer to the question that they are trying to answer.[226] Asked to solve a murder mystery, a deliberating group did far better when its members were told that they had sufficient clues to "determine" the identity of the guilty suspect than when they were told to decide which suspect was "most likely to have committed the crime."[227] Hence "adequate

---

220 *See* James S. Fishkin & Robert C. Luskin, *Bringing Deliberation to the Democratic Dialogue*, *in* THE POLL WITH A HUMAN FACE 3, 29–31 (Maxwell McCombs & Amy Reynolds eds., 1999) (suggesting that deliberation can produce accurate judgments).

221 *See* Gigone & Hastie, *Proper Analysis*, *supra* note 1, at 165.

222 Stasser & Dietz-Uhler, *supra* note 86, at 40.

223 *See* Norbert L. Kerr & Ernest S. Park, *Group Performance in Collaborative and Social Dilemma Tasks: Progress and Prospects*, *in* GROUP PROCESSES, *supra* note 86, at 107, 110 (describing group successes in processing large amounts of information).

224 *See* Hastie, *supra* note 86, at 148.

225 *Id.*

226 *See* Garold Stasser & Dennis Stewart, *Discovery of Hidden Profiles by Decision-Making Groups: Solving a Problem Versus Making a Judgment*, 63 J. PERSONALITY & SOC. PSYCHOL. 426, 432–33 (1992).

227 *Id.* at 428.

Imaged with Permission from N.Y.U. Law Review

consideration of unshared, critical information during group discussion" appears to be affected by "how members construe their decision-making task";[228] those who believe that they are solving a problem with a correct solution are more likely to explore shared information than those who think that they are reaching a consensus. It follows that "discussions may be more data driven and less consensus driven when members believe that a demonstrably correct answer exists."[229] Even here, however, the member who knows the right solution usually requires some initial support in the group; otherwise the group will frequently fail.[230]

Another study finds that groups performed exceedingly well, far better than individual members, in two complex tasks that had demonstrably correct solutions.[231] The first involved a statistical problem, requiring subjects to guess the composition of an urn containing blue balls and red balls. The second involved a problem in monetary policy, asking participants to manipulate the interest rate to steer the economy in good directions. People were asked to perform as individuals and in groups. The basic results for the two experiments were similar. Groups significantly outperformed individuals. On a scale of 1 to 100, the average group score in the urn test was 86.8, as opposed to 83.7 for individuals—a highly significant difference statistically. For the monetary policy problem, the difference was essentially identical. Interestingly, groups did not, on balance, take longer to make a decision. In terms of both accuracy and time, there were no differences between group decisions made with a unanimity requirement and group decisions made by majority rule.

How can these results be explained? An obvious possibility is that group processes play a small role and that the group's discussion is simply the average of individual judgments. On this view, the judgments of these deliberating groups simply *were* statistical judgments. But the evidence is inconsistent with this hypothesis; groups in these cases did far better than their average member.[232] Even more remarkably, the performance of the median player did not explain the performance of the group. An alternative hypothesis is that each group contained one or more strong analysts, who were able to move

228 *Id.* at 432.

229 *Id.* at 433.

230 *See* MacCoun, *supra* note 10, at 120 (showing that groups will not arrive at accurate answer unless that answer begins with significant support).

231 ALAN S. BLINDER & JOHN MORGAN, ARE TWO HEADS BETTER THAN ONE? AN EXPERIMENTAL ANALYSIS OF GROUP VS. INDIVIDUAL DECISIONMAKING 1, 6, 15, 46–47 (Nat'l Bureau of Econ. Research, Working Paper No. 7909, 2000).

232 *Id.* at 41.

Imaged with Permission from N.Y.U. Law Review

the group in the right direction. But in the experiments, there is little support for this hypothesis either. "In the end, we are left to conclude that neither the average player, nor the median player, nor the best player determine the decisions of the group."[233] It seems that, in these experiments, the better decisions by groups resulted from the fact that the best points and arguments turned out to spread among the various individual players. Here we find some basis for the claim that, under appropriate conditions, groups can do much better than individuals. The relevant conditions appear to include highly competent group members attempting to solve statistical problems that all members knew to have demonstrably correct answers.

2. *The Deliberative Opinion Poll*

In an important combination of theoretical and empirical work, James Fishkin has pioneered the idea of a "deliberative opinion poll." In deliberative opinion polls, small groups, consisting of highly diverse individuals, are asked to come together and to deliberate about various issues.[234] Fishkin has conducted deliberative opinion polls in several nations, including the United States, England, and Australia. Fishkin finds some noteworthy shifts in individual views, and he evidently believes that the deliberative process produces learning and hence improvements in people's judgments.[235] Because of the nature of the deliberative opinion poll, it is not possible to test for the amplification of errors, hidden profiles, or cascade effects. Nor is it possible, at least in much of the existing work, to see whether deliberating groups have moved toward correct answers—a point to which I shall return. But Fishkin does not find a systematic tendency toward group polarization. In his studies, individuals shift both toward and away from the median of predeliberation views.[236] It is therefore tempting to conclude that properly structured deliberation can avoid some or possibly even all of the problems traced here.

In England, for example, deliberation led to reduced interest in using imprisonment as a tool for combating crime.[237] The percentage believing that "send[ing] more offenders to prison" is an effective way to prevent crime went down from 57% to 38%; the percentage believing that fewer people should be sent to prison increased from 29% to 44%; belief in the effectiveness of "stiffer sentences" was

---

233 *Id.* at 46.

234 *See* FISHKIN, *supra* note 5, at 161–86 (discussing deliberative polling).

235 *See id.* at 168 (stating that voters who participated in poll "changed in dramatic and coherent ways").

236 *Id.* at 167–68, 177–81.

237 *Id.* at 167–68.

Imaged with Permission from N.Y.U. Law Review

reduced from 78% to 65%.[238] Similar shifts were shown in the direction of greater enthusiasm for procedural rights of defendants and increased willingness to explore alternatives to prison.[239] In other experiments with the deliberative opinion poll, shifts included a mixture of findings, with larger percentages of individuals concluding that legal pressures should be increased on fathers for child support (from 70% to 85%) and that welfare and health care should be turned over to the states (from 56% to 66%).[240]

On many particular issues, including the two just mentioned, the effect of deliberation was to create an increase in the popularity of the view that initially had majority support within the group.[241] These findings are consistent with the prediction of group polarization. But this was hardly a uniform pattern. On some questions, deliberation increased the percentage of people holding a minority position (with, for example, a jump from 36% to 57% of people favoring policies making divorce "harder to get").[242] These are not the changes that would be predicted by group polarization.

We should be careful, however, about celebrating the results found in Fishkin's studies, at least as I have described them thus far. What would be most revealing would be a series of findings in which deliberative opinion polls, using the same methods, led people to *correct* answers on questions with objectively correct answers. If there were evidence that this happened, and happened systematically, then we would have reason to be confident that deliberation is actually doing a lot of good. Suppose, for example, that we can agree that doubling the minimum wage would have significant disemployment effects, or that a significant increase in capital punishment would have at least some deterrent effect on crime, or that global warming is likely to occur, or that DDT has significant effects in reducing the incidence of malaria in poor countries. Suppose too that diverse groups, consisting of people with widely varying views on these issues, were to assemble for a deliberative opinion poll. If members moved systematically toward the correct answers, then deliberation would, by hypothesis, be working. But I do not believe that Fishkin's findings

238 *Id.* at 178–79.

239 *Id.* at 179–80.

240 Fishkin & Luskin, *supra* note 220, at 23.

241 *See id.* at 22 (showing jump, on scale of 1 to 4, from 3.51 to 3.58 in intensity of commitment to reducing deficit; showing jump, on scale of 1 to 3, from 2.71 to 2.85 in intensity of support for greater spending on education; showing jump, on scale of 1 to 3, from 1.95 to 2.16, in commitment to aiding American business interests abroad).

242 *Id.* at 22–23 (showing an increase, on scale of 1 to 3, from 1.40 to 1.59 in commitment to spending on foreign aid; also showing decrease, on scale of 1 to 3, from 2.38 to 2.27 in commitment to spending on social security).

Imaged with Permission from N.Y.U. Law Review

provide a rigorous test of whether group members are moving in the right directions. Instead, they show only that people both learn and move—an impressive finding, to be sure, but one that does not show that they move systematically toward correct answers.

Why doesn't the deliberative opinion poll always produce group polarization? There are several possible answers. First, and probably most important, Fishkin's studies presented participants with a set of written materials that attempted to be balanced and that contained detailed arguments supporting both sides. The likely consequence would be to move people in different directions from those that would be expected by simple group discussion, unaffected by external materials inevitably containing a degree of authority. Indeed, the very effort to produce balance introduces new elements into group deliberations, simply because the argument pool is different from what it would be if all claims were generated independently by group members. Second, Fishkin's deliberators did not vote as a group. While group polarization is observed when no group decision is expected, the extent of polarization is likely to decrease, simply because members have not been asked to sign onto a group decision as such. Third, Fishkin's groups were overseen by a moderator, concerned to ensure a level of openness and likely to alter some of the dynamics discussed here. A moderator, even a neutral one, can do a great deal to reduce polarization, by altering both informational and reputational influences.

Some people are optimistic about the results of deliberative opinion polls and want them to be used more broadly.[243] I agree that they show a great deal of promise, certainly as compared to polls that consist of mere snapshots of unreflective opinions. But without systematic evidence that people have been moving in better directions, we should be cautious about the current findings. From existing deliberative opinion polls, taken together with other evidence about group processes,[244] it is not at all clear whether deliberation will increase accuracy, even under Fishkin's conditions.

The deliberative opinion poll does, however, provide important lessons about appropriate institutional design for deliberating bodies. Group polarization can be heightened, diminished, and possibly even eliminated with seemingly small alterations in institutional arrangements. To the extent that informational pressure and social influences are likely to have unfortunate effects, valuable correctives can be

[243] *See* ACKERMAN & FISHKIN, *supra* note 6, at 3 (proposing national day for political deliberation).

[244] *See supra* notes 129–44 and accompanying text (discussing amplification of errors).

Imaged with Permission from N.Y.U. Law Review

introduced, perhaps above all by exposing group members, at one point or another, to arguments to which they are not antecedently inclined. Let us now approach this question more systematically.

III
REMEDIES AND REFORMS

How might group deliberation be improved? How can groups counteract the problems I have emphasized? An understanding of problems with deliberating groups helps to specify promising remedies, increasing the likelihood that deliberation will lead to more accurate or sensible solutions. These remedies hold out considerable promise for ensuring that deliberation does what it is intended to do.

If, for example, mistakes come from informational and reputational pressure, then the solution is to take steps to increase the likelihood that people will disclose what they know. If people are asked to think critically rather than to join the group, and they are told that the group seeks and needs individual contributions, then disclosure is more likely. Consider a redefinition of what it means to be a "team player." Frequently a team player is thought to be someone who does not upset the consensus; but it would be possible to understand team players as those who increase the likelihood that the team will be right—if necessary, by disrupting the conventional wisdom.

Institutional reforms can do a great deal to counteract the problems caused by informational and reputational pressure. The most difficult problem is the propagation of error. If group members use the availability heuristic, or if they fall prey to optimistic bias, blunders will result unless they are corrected by one or more group members. Even here, however, the best solution is to attempt to ensure that group members say what they believe to be true.

For those who seek to diminish the effects of informational pressure and social influences, a cautionary note remains. We can imagine groups that actually benefit from these effects, and hence from cascades and polarization.[245] Sometimes it is good for people to silence themselves; sometimes their contributions would be unhelpful, because what they believe that they know is false.[246] Suppose that some group members have a terrible idea about how to stabilize the economy, litigate a case, or reduce the threat of terrorism. If so, we should be grateful for informational pressure and social influences

[245] Hidden profiles, of course, are never desirable, assuming the truth of the information that is hidden.

[246] Of course falsity can sometimes contribute to truth; but it frequently does not.

Imaged with Permission from N.Y.U. Law Review

that make them defer to those who know much better. As a result of those pressures, the group's decisions will be improved.

We have seen that polarization might lead people in the right direction; the question is whether a more extreme version of members' antecedent tendency is correct, and that question must be answered on its merits. The process of polarization does not provide that answer. Or consider a cascade in which the early movers actually know the truth, and those who follow them are ignoring private information that they believe to be true but that would, on reflection, turn out to be erroneous or misleading. If so, the followers are not only rational in disregarding what they know; they also lead the group in a better direction because they do not give it bad signals. Those who participate in cascades are acting rationally; but the more important point is that if those who start cascades are correct, both individuals and groups are better off as a result. The only problem—and it is a serious one—is that many cascade participants will fail to disclose accurate information, and for that reason the group will suffer. The discussion in this section offers some lessons for how to reduce the risk of erroneous cascades, by diminishing the effects of informational and reputational pressures.

Let us focus, then, on the standard cases in which deliberating groups will do worse if they do not learn what group members know. For private and public institutions, the overriding question is how to alter people's incentives in such a way as to increase the likelihood of disclosure. Many possibilities might be imagined here. Consider two experiments that have more general implications.

A. Restructured Incentives

Is it possible to reduce the pressures that lead group members to silence themselves? Is it possible to ensure that people will internalize some of the benefits that accrue to the group from disclosure?

1. Overcoming Reputational Influences: Priming Critical Thinking

Self-silencing is partly a product of social norms—of a sense that people will be punished rather than rewarded for disclosing information that departs from the group's inclination. Group processes can aggravate or eliminate this effect. If consensus is prized, and known to be prized, then self-silencing will be more likely. If the group is known to welcome new and competing information, then the reward structure will be fundamentally different.

Evidence for this claim comes from hidden profile experiments that "primed" people by asking them to engage in a prior task that

Imaged with Permission from N.Y.U. Law Review

involved either "getting along" or "critical thinking."[247] Primed by a task that called for critical thinking, people were far more likely to disclose what they knew, and there was a quite substantial reduction of hidden profiles.[248] (Team players think critically, and they do not always get along.) For both private and public groups, the general lesson is clear. If norms favor disclosure of privately held information, then self-silencing will be reduced significantly; deliberation is likely to benefit as a result. Social norms and institutional culture can go a long way toward reducing the effects of social pressures.

2. *Overcoming Informational Influences: Rewarding Group Success*

We have seen that people often do not disclose what they know because they receive only a fraction of the benefits of disclosure; this problem is compounded if private information seems likely to be erroneous in light of what others have said. But how would groups perform if individuals knew that they would be rewarded, not if their own answer was correct, but if the majority of the group was correct? It might be speculated that in a situation of this kind, hidden profiles, cascades, and group polarization would be reduced dramatically. The reason is that when people are rewarded when their group is right, they are far more likely to reveal, to that group, what they actually know. In such a situation, incentives are restructured so that people internalize the benefits of disclosure.

For supportive evidence, consider an intriguing variation on the urn experiment, where subjects were paid $2 for a correct *group* decision and penalized $2 for an incorrect *group* decision, with the group decision determined by majority rule.[249] People were neither rewarded nor punished for a correct individual decision. The result was that, in 92% of cases, people's announcement matched their private draw.[250] And because people revealed their private signals, the system of majority rule produced a huge increase in fully informed decisions—that is, the outcomes that someone would reach if he was somehow able to see all private information held by group members. As an example, consider this round from the majority rule experiment[251] (the actual urn was A):

---

247 Stasser & Titus, *Hidden Profiles*, *supra* note 10, at 309.

248 *Id.* at 309–12.

249 Hung & Plott, *supra* note 192, at 1511.

250 *Id.* at 1517–18.

251 *Id.* at 1515.

Imaged with Permission from N.Y.U. Law Review

TABLE 3: NO CASCADE

	1	2	3	4	5	6	7	8	9
Private Draw	A	A	A	A	B	A	A	A	B
Decision	A	A	A	A	B	A	A	A	B

What is the explanation for this significantly reduced level of cascades in a system of majority rule? The answer lies in the fact that the individual knows that he has nothing to gain from a correct *individual* decision and everything to gain from a correct *group* decision. As a result, it is in the individual's interest to say exactly what he sees because it is the accurate announcement, from each person, that is most likely to promote an accurate group decision. A simple way to understand this point is to assume that a group has a large number of members and that each member makes an announcement that matches his private draw. As a statistical matter, it is overwhelmingly likely that the majority's position will be correct.

Though this experiment is highly artificial, an emphasis on the importance of group success should improve decisions in many real-world contexts. Suppose, for example, that members of a jury are strongly committed to ensuring an accurate outcome from the group; if so, they will be more likely to disclose what they know. Whistleblowing is often a product, not of the whistleblower's narrow self-interest (which may well argue in favor of self-silencing), but of a belief that it is important to take steps to ensure that the organization or group acts properly. The general lesson is that identification with the group's success is more likely to ensure that people will say what they know. And if group members focus on their own prospects and accuracy, rather than that of the group, the group is more likely to err. Both social norms and material incentives play crucial roles in establishing the priorities, on this count, of group members.

B. Devil's Advocates

How can institutional design take advantage of these findings? If hidden profiles and self-silencing are the source of group failure, then an obvious response is to ask some group members to act as "devil's advocates," urging a position that is contrary to the group's inclination.[252] This was a central suggestion of both the Senate Committee reporting on intelligence failures in connection with Iraq and of the review board that investigated large blunders at NASA.[253]

---

[252] *See* JANIS, *supra* note 11, at 267.

[253] *See supra* notes 12–20.

Imaged with Permission from N.Y.U. Law Review

Those assuming the role of devil's advocate will not face the reputational pressure that comes from rejecting the dominant position within the group; they have been charged with doing precisely that. And because they are asked to take a contrary position, they are freed from the informational influences that can lead to self-silencing. Hidden profiles are less likely to remain hidden if one or more group members are told to disclose the information they have, even if that information runs contrary to the apparent tendency within the group. In at least one well-known case, this approach appeared to work. "During the Cuban missile crisis, President Kennedy gave his brother, the Attorney General, the unambiguous mission of playing devil's advocate, with seemingly excellent results in breaking up a premature consensus"[254]—a consensus that might well have led to war.

Unfortunately, research on devil's advocacy in small groups does not provide conclusive evidence of the effectiveness of devil's advocacy in real-world settings.[255] To be sure, many experimenters have found that protection of genuine dissenting views can enhance group performance.[256] But a formal requirement of devil's advocacy enhances group performance far less than does the articulation of genuine dissent. When an advocate's challenges to a group consensus are insincere, members discount his arguments accordingly. At best, he merely facilitates a "multisided examination of the problem at hand."[257] Because devil's advocates have no incentive to sway the group's members to their side, they accomplish their task even if they allow the consensus view to refute the unpopular dissenting arguments. Unlike a genuine dissenter, the devil's advocate has little to gain by zealously challenging the dominant view—and as a result tends not to persist in challenging the consensus.[258]

In any case, the perceived sincerity of a dissenter is an important factor in determining minority influence.[259] An insincere devil's advocate is unlikely to provide much help. The lesson is that if devil's advocacy is to work, it is because the group attempts to ensure that the dissenter actually means what he is saying. If so, better decisions can be expected.

---

254 JANIS, *supra* note 11, at 268.

255 Gary Katzenstein, *The Debate on Structured Debate: Toward a Unified Theory*, 66 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 316, 317–18 (1996).

256 Alexander L. George & Eric K. Stern, *Harnessing Conflict in Foreign Policy Making: From Devil's to Multiple Advocacy*, 32 PRESIDENTIAL STUD. Q. 484, 486 (2002).

257 *Id.*

258 *Id.*

259 Serge Moscovici, *Social Influence and Conformity*, *in* 2 HANDBOOK OF SOCIAL PSYCHOLOGY 347, 359–65 (Gardner Lindzey & Elliot Aronson eds., 3d ed. 1985).

Imaged with Permission from N.Y.U. Law Review

C. *Enlisting High-Status Contrarians and Leadership*

Some people are more likely to silence themselves than others. For example, group members are less likely to conform if they have high social status or are extremely confident about their own views.[260] In a complementary finding, members of low status groups—less educated people, African Americans, sometimes women—have been shown to carry less influence within deliberating groups than members with higher status.[261] On juries, lower status members, as measured by their occupations and sex, have been found to be less active and less influential in deliberation.[262] Creative groups would do well to take account of these findings.

For example, the problem of unshared information is reduced when that information is held by a leader within a group; not surprisingly, the leader's words count, because people listen to what leaders have to say.[263] In one experiment, a medical team consisting of a resident physician, an intern, and a third-year medical student showed a tendency to emphasize unshared items stressed by the *resident*—and in this respect did not fall prey to the problem of hidden profiles.[264] More generally, those experienced in the task at hand are more likely to mention and to repeat unshared information.[265]

One reason for these findings is that those with higher status or competence are less subject to the reputational pressures that come from emphasizing unshared information.[266] Another reason is that leaders and experts are more likely to think that their own information is accurate and worth disclosing to the group, notwithstanding the fact that the information held by other group members cuts in the other direction.

The simplest lesson is that leaders and high-status members can do groups a great service by asserting a contrary view, at least for purposes of argument.[267] In a similar vein, group leaders should be reluctant to state a firm view at the outset and should, in that way, allow space for more information to emerge.

---

260 *See* Christensen & Abbott, *supra* note 112, at 272–76.

261 *Id.*

262 *See* Ridgeway, *supra* note 183, at 54.

263 Stasser & Titus, *Hidden Profiles*, *supra* note 10, at 308.

264 *Id.*

265 *Id.*

266 *Id.*

267 *Cf.* JANIS, *supra* note 11, at 262–63 (emphasizing need for leaders to be willing to accept criticism of their own judgments).

Imaged with Permission from N.Y.U. Law Review

D. Predeliberation Anonymity, Secret Ballots, and the Delphi Method

To overcome social influences, people might be asked to register their opinions anonymously, either in advance of deliberation or after it has occurred. The secret ballot can be understood as an effort to insulate people from reputational pressures and to permit them to say what they believe.[268] Many institutions should consider more use of the secret ballot simply to elicit more information.

As an ambitious effort to implement this idea, consider the Delphi Technique, a process for aggregating the views of group members. The Delphi Technique has four key features.[269] First, it ensures the anonymity of all members through a self-administered questionnaire. The purpose of anonymity is precisely "to diminish the effects of social pressures, as from dominant or dogmatic individuals, or from a majority."[270] Second, the technique is iterated, and there is a system for controlled feedback on the judgments of others. Members make individual estimates; all members are informed of the views of other members; and there are additional rounds of estimates, allowing feedback until there is a desired level of convergence. Third, group members are permitted to communicate, but sometimes only their ultimate conclusions (generally in the form of summary statistics involving quartiles or ranges); and typically the conclusions, given anonymously, are provided to others by a facilitator or monitor team, often in the form of a simple summary such as a mean or median value of the group response. Thus "the feedback comprises the opinions and judgments of all group members and not just the most vocal."[271] (Note here that the Delphi Technique is more successful when group members are provided not only with the mean or median estimate, but also with reasons given by group members for their views.[272] An account of reasons is most likely to move people in the direction of the correct answer.[273]) Fourth, and finally, the judgments of group members are subject to a statistical aggregation.

The Delphi Technique provides a sharp contrast with efforts to obtain the judgments of statistical groups and also with interacting groups containing open deliberation. In several contexts, the Delphi

---

268 *See* TIMUR KURAN, PRIVATE TRUTHS, PUBLIC LIES: THE SOCIAL CONSEQUENCES OF PREFERENCE FALSIFICATION 13–15 (1995) (discussing secret ballot as protection against social pressures).

269 *See* Rowe & Wright, *supra* note 106.

270 *Id.* at 126.

271 *Id.*

272 *Id.* at 129.

273 *Id.* at 129–30.

Imaged with Permission from N.Y.U. Law Review

Technique has provided more accuracy than open discussion.[274] For general almanac questions, the Delphi Technique was found to produce better answers than individual estimates, though open discussion did still better, apparently because it served to correct errors.[275] A natural alternative to the Delphi Technique would be a system in which ultimate judgments were stated anonymously, but only after deliberation. Anonymity would insulate group members from reputational pressure, and to that extent could reduce the problem of self-silencing. But it would do little to reduce informational pressure.

E. Roles, Experts, and Forewarning

Imagine a deliberating group consisting of people with specific roles appreciated and known by all group members. One person might be understood to have medical expertise; a second might be a lawyer; a third might know about public relations; a fourth might be a statistician. In such a group, it might be hypothesized that sensible information aggregation would be far more likely, simply because each member knows that each of the others has something particular to contribute. Hidden profiles should be less likely to remain hidden if there is a strict division of labor, in which each person is knowledgeable, and known to be knowledgeable, about something in particular.[276]

Several experiments support the hypothesis.[277] In one such experiment, each member of a three-person group was given a good deal of information about one of three murder suspects.[278] In half of these groups, the "expertise" of each member was publicly identified to all before discussion began; in half of them, there was no such public identification of experts. The bias in favor of shared information was substantially reduced in those groups in which experts were publicly identified as such.[279] The reduction of the bias was significantly smaller when there was no *public* identification of experts and

---

274 *See id.* at 130; Hastie, *supra* note 86, at 139.

275 *See* Hastie, *supra* note 86, at 139–45.

276 *See* Garold Stasser, *The Uncertain Role of Unshared Information in Collective Choice, in* SHARED COGNITION IN ORGANIZATION: THE MANAGEMENT OF KNOWLEDGE 49, 56–57 (Leigh L. Thompson et al. eds., 1999).

277 *See* Stasser & Titus, *Hidden Profiles*, *supra* note 10, at 308 (citing studies showing that "when the bearer of unique information was labeled an expert, the group seemingly paid more attention to the information"); Garold Stasser et al., *Expert Roles and Information Exchange During Discussion: The Importance of Knowing Who Knows What*, 31 J. EXPERIMENTAL SOC. PSYCHOL. 244, 248–49, 256 (1995) (showing that assigning expert roles led to more discussion of unshared data).

278 *See* Stasser et al., *supra* note 277, at 248–49.

279 *Id.* at 249, 259–62.

Imaged with Permission from N.Y.U. Law Review

when each group member was simply told, by the experimenter, that he or she was an expert on a particular candidate.[280] The lesson is clear: If a group seeks to obtain the information that its members hold, it would make sense to inform all group members, before deliberation begins, that different members have different, and relevant, information to contribute. Unfortunately, however, the effect of role assignment in reducing hidden profiles is not huge.[281]

F. General Lessons

These various findings offer general lessons about how deliberating groups might significantly reduce the adverse effects of informational influences and social pressures. The lessons apply to such diverse groups as corporate boards, juries, multimember judicial panels, and administrative agencies.[282] If information is dispersed within the group, leaders would do well not to state a firm view at the outset; they might well refrain from expressing any opinion at all until other people have said what they think. Following the model of Franklin Delano Roosevelt, they might indicate sympathy for a wide range of views, encouraging diverse opinions to arise.[283] They might suggest in particular that they welcome information and perspectives that diverge from their own. A degree of impartiality, on the part of leaders, would go a long way toward encouraging diversity of views. And if reasonable alternatives are not being discussed, group members might be assigned the task of developing and presenting them. Independent subcommittees might be asked to generate new views, possibly views that compete with one another.

Consider, for example, the CIA and NASA examples with which I began. Subsequent investigators found that both agencies contained enough information to prevent their large-scale blunders. If internal processes had been properly structured, those blunders would have been less likely to occur. Suppose that a norm of critical thinking had been encouraged, so that employees would have felt free to challenge assumptions about weapons of mass destruction in Iraq or about the assumed safety of the Challenger and Columbia flights. Or suppose that both agencies had created an internal system of checks and balances, ensuring careful attention to competing views. If so, it would have been far more likely that relevant information would have emerged, and been taken seriously, during internal processes.

---

280 *Id.* at 262.

281 Stasser & Titus, *Hidden Profiles*, *supra* note 10, at 310 (summarizing studies).

282 For overlapping prescriptions, see JANIS, *supra* note 11, at 262–73.

283 *See* ROBERT H. JACKSON, THAT MAN: AN INSIDER'S PORTRAIT OF FRANKLIN D. ROOSEVELT 115–17 (John Q. Barrett ed., 2003).

Imaged with Permission from N.Y.U. Law Review

Of course time is limited, and prescriptions that are suitable for some organizations will not be suitable for others. In the context of jury deliberations, subcommittees would make little sense; what is required is an initial degree of openness in which jurors explore relevant facts before announcing a conclusion. For regulatory agencies, by contrast, competing subdivisions can help to ensure a range of perspectives. In this vein, Christopher Edley has suggested that Congress should create, within the Department of Homeland Security, an independent Office on Rights and Liberties, whose specific mission would be to ensure that the effort to protect the nation from terrorist threats does not unduly compromise individual rights.[284] In Edley's account, the Office would receive and address public complaints about rights violations; it would also make classified quarterly reports to Congress and the President, along with unclassified reports to the public.[285] The proposal deserves serious consideration as a check on amplification of errors, hidden profiles, and group polarization.[286]

An optimistic view of the structure of the Environmental Protection Agency would suggest that the proliferation of offices with overlapping tasks—including a pro-regulatory Air Office and a more technocratic Planning Office—ensures a kind of internal system of checks and balances.[287] Under existing law, the independent regulatory agencies, including the Securities and Exchange Commission and the Federal Communications Commission, may not have more than a bare majority of their members from a single political party.[288] This limitation might be understood as an effort to protect against the deliberative pathologies that are likely to result if deliberations are restricted to like-minded people.

Many variations on these themes might be imagined. My goal here has not been to set out an institutional blueprint, but to suggest some general points that deliberating groups might take into account when structuring their processes for eliciting and aggregating information and points of view.

---

284 *See* Christopher Edley, *A U.S. Watchdog for Civil Liberties*, WASH. POST., July 14, 2002, at B7.

285 *See id.*

286 The 9/11 Commission, citing the lack of a "voice within the executive branch" designed to consider liberty concerns, has made a similar recommendation. *See* NAT'L COMM'N ON TERRORIST ATTACKS UPON THE U.S., THE 9/11 COMMISSION REPORT 395 (2004), *available at* http://www.gpoaccess.gov/911/index.html.

287 *See* BRUCE A. ACKERMAN & WILLIAM T. HASSLER, CLEAN COAL/DIRTY AIR 79–86 (1981) (describing roles of respective offices).

288 *See, e.g.*, 15 U.S.C. § 78d (1994); 47 U.S.C. § 154(b)(5) (2000).

Imaged with Permission from N.Y.U. Law Review

IV
INFORMATION MARKETS

Deliberation is one way to aggregate privately held information; but there are many other possibilities. Open-source software, for example, provides a method by which decentralized "bits" of privately held information can be drawn together in software design, thus ensuring improvements that go far beyond the capacities of small groups of experts.[289] With open-source software, expert groups do not deliberate about technological improvements; instead numerous contributors can bring their creativity and knowledge to bear.

More generally, the Internet itself is easily used as an aggregative mechanism.[290] For example, a "wiki" is a website that allows any user to add material and to edit what previous users have done.[291] Wikipedia operates as a free, web-based encyclopedia[292] that attempts to take advantage of the information held by thousands of contributors ("Wikipedians"), who add to and edit the encyclopedia. We should see wikis through the lens of Hayek's emphasis on the highly dispersed nature of information in society and the value of developing mechanisms that serve to aggregate that information.[293]

In a similar vein, a great deal of recent attention has been paid to weblogs, which can serve to elicit and aggregate the information held by countless contributors.[294] If thousands of people are maintaining their own "blogs," they should be able to act as fact-checkers, and as supplemental information sources, for the more prominent members of the mass media.[295] And if tens of thousands of people are reading the most prominent blogs, then errors, on the part of bloggers, should

---

289 *See* LAWRENCE LESSIG, FREE CULTURE 44 (2004) (emphasizing virtues of open-source software).

290 For a superb general discussion of aggregation through peer-to-peer interactions with particular reference to the Internet, see Yochai Benkler, *Coase's Penguin, or, Linux and* The Nature of the Firm, 112 YALE L.J. 369 (2002).

291 *Wiki*, *in* WIKIPEDIA: THE FREE ENCYCLOPEDIA, *at* http://en.wikipedia.org/wiki/Wiki Wiki (last modified Apr. 3, 2005).

292 *See Wikipedia*, *in* WIKIPEDIA: THE FREE ENCYCLOPEDIA, *supra* note 291. I am grateful to Agata Waclawik for this reference.

293 *See generally* Hayek, *supra* note 128. Note in this regard that Lawrence Lessig plans to update his book CODE AND OTHER LAWS OF CYBERSPACE (2000) by posting it as a wiki. *See* Lawrence Lessig, *Tis the Season: II*, LESSIG BLOG (posted Dec. 24, 2004, 1:51 PM), *at* http://www.lessig.org/blog/archives/002358.shtml.

294 For an overview, see Daniel W. Drezner & Henry Farrell, The Power and Politics of Blogs (July 2004) (unpublished manuscript, on file with the *New York University Law Review*).

295 For a popular overview, see generally HUGH HEWITT, BLOG: UNDERSTANDING THE INFORMATION REFORMATION THAT'S CHANGING YOUR WORLD (2005).

Imaged with Permission from N.Y.U. Law Review

be corrected quickly.[296] Of course blogs may (and do) also suffer from amplification of error, hidden profiles, cascade effects, and group polarization. But they hold out the promise of aggregating information held by large numbers of people. Sheer numbers are playing a large role here, because information aggregation is likely to work best when many people are involved; but it is also important that reasons and information are being exchanged, in a way that can lead to corrections and real creativity.

Another way to aggregate information is to rely on the price signal, which has a similar aggregative function. Consider a familiar informal challenge when people disagree on some question of fact: "Want to bet?"[297] The point of the challenge is to suggest that the speaker is quite confident of her judgment, enough so to ask the person with whom she disagrees to back her conviction with dollars. Not infrequently, the challenge is successful in the sense that it operates to establish, to all concerned, that one or another belief is weakly held. But it is possible to use economic incentives far more formally and systematically.

In fact the great advantage of the price signal is that it aggregates both the information and the tastes of numerous people, producing judgments that incorporate more material than could possibly be assembled by any central planner, even one who insists on deliberation with and among experts. Recall Hayek's claim about the price system and its aggregative properties,[298] a claim that I have used as the basis for a Hayekian challenge to generalized celebrations of deliberation. We can even see a link between the Hayekian claim and the Condorcet Jury Theorem; precisely because many people are making purchasing decisions, their aggregate judgments are highly

---

296 Ana Marie Cox ("Wonkette"), Presentation at the American Political Science Association (Sept. 3, 2004). For discussion of the role of bloggers in ferreting out errors in the mass media, see HEWITT, *supra* note 295.

297 See the outline of the debate between Julian Simon and Paul Ehrlich, *infra* note 390.

298 *See* Hayek, *supra* note 128, at 524–28 (discussing dispersed nature of knowledge and value of aggregating information through price system). Information markets (sometimes called prediction markets) have received discussion in the literature. *See generally* Joyce Berg et al., *Results from a Dozen Years of Election Futures Markets Research*, *in* HANDBOOK OF RESULTS IN EXPERIMENTAL ECONOMICS (Charles R. Plott & Vernon L. Smith eds., forthcoming 2005) (exploring results of Iowa markets), *available at* http://www.biz.uiowa.edu/faculty/trietz/papers/iemresults.pdf (working draft, March 2003); Joyce Berg et al., *What Makes Markets Predict Well? Evidence from the Iowa Electronic Markets*, *in* UNDERSTANDING STRATEGIC INTERACTION 444 (Wulf Albers et al. eds., 1997) (discussing Iowa Electronic Markets); Robert Forsythe et al., *Anatomy of an Experimental Political Stock Market*, 82 AM. ECON. REV. 1142 (1992), (discussing origins and performance of Iowa markets); Robert Forsythe et al., *Wishes, Expectations, and Actions: A Survey on Price Formation in Election Stock Markets*, 39 J. ECON. BEHAV. & ORG. 83 (1999) [hereinafter Forsythe et al., *Wishes*] (reviewing field experiments on information markets).

Imaged with Permission from N.Y.U. Law Review

likely to be right, at least if most purchasers have relevant information. And if an emphasis is placed on the information-aggregating properties of markets, it would seem plain that if we are attempting to improve on the answer produced by statistical means and deliberating groups, we might consider an increasingly popular possibility: *Create a market.*[299] Information markets, a recent innovation, have proved remarkably successful at forecasting future events; they seem to do far better, in many domains, than deliberating groups. Such markets are worth sustained attention, in part because they offer important lessons about how to make deliberation go better or worse, and in part because they provide a useful model for many private and public organizations.

A central advantage of information markets is that they impose the right incentives for people to disclose the information that they hold. Recall that in a deliberating group, members often have little incentive to say what they know. By speaking out, they provide benefits to others, while possibly facing high private costs. Information markets realign incentives in a way that is precisely designed to overcome these problems. Because investments in such markets are generally not disclosed to the public, investors need not fear reputational sanctions if, for example, they have predicted that a company's sales will be low or that a certain candidate will be elected president. And because people stand to gain or lose from their investments, they have a strong incentive to use (and in that sense to disclose) whatever private information they hold; they can capture, rather than give to others, the benefits of disclosure. The use of private information will be reflected in the price signal. In these crucial ways, the problems that infect deliberating groups are largely eliminated in information markets. I have outlined a series of initiatives that should improve the performance of deliberating groups. But information markets automatically do much of the work of those initiatives.

We have seen that optimal deliberation is structured in a way that permits relevant and correct information to emerge—and that reduces the likelihood that useless, biased, or incorrect information will undermine deliberation. For their part, information markets impose strong incentives for traders to ferret out accurate information. Perhaps

---

[299] *See generally* Justin Wolfers & Eric Zitzewitz, *Prediction Markets*, 18 J. ECON. PERSP. 107 (2004) (valuable overview of prediction markets); Michael Abramowicz, *Information Markets, Administrative Decisionmaking, and Predictive Cost-Benefit Analysis*, 71 U. CHI. L. REV. 933 (2004) (recommending use of information markets by administrative agencies); Saul Levmore, *Simply Efficient Markets and the Role of Regulation: Lessons from the Iowa Electronic Markets and the Hollywood Stock Exchange*, 28 J. CORP. L. 589 (2003) (emphasizing accuracy of information markets, even when there are relatively few traders).

Imaged with Permission from N.Y.U. Law Review

most importantly, information markets have been found not to amplify individual errors but to eliminate them; the prices that result from trading prove reliable even if many individual traders err.[300] Traders do not trade blindly; they are entirely able to stop trading, for a moment or much more, in order to retrieve better information that will give them an advantage. But in some deliberating groups, participants cannot leave; they must continue deliberating, and the necessary information is, at best, dispersed and locked within individual participants. Well-functioning systems of deliberation encourage group members to act dynamically to acquire further information, just as markets tend to do.

Of course investors, like everyone else, are subject to the informational pressure imposed by the views of others. But a market creates strong incentives for revelation of whatever information people actually hold. For small groups, of course, information markets are likely to be too "thin" to be useful; a certain number of investors is required to get a market off the ground.[301] Hence feasibility is a serious constraint on the use of information markets. In some contexts, however, private and public organizations might use markets as a complement to or even a substitute for deliberation.

How might information markets be used? Consider a few possibilities:

1. Uncertain about sales projections for the future, a company might not ask its employees to make predictions or to deliberate with one another. Instead it might create an information market, in which employees are allowed to place anonymous bets about likely outcomes.[302]
2. As an aid to its assessment of future events in the world, the White House and the Department of Defense might maintain an information market in which investors predict outcomes of national importance—for example, the likelihood that the government of Iran will be toppled in the next calendar year, that there will be a terrorist attack in Europe within the same period, or that free elections will be held in Iraq by a specified time.

---

300 *See infra* notes 333–40.

301 *But see* Levmore, *supra* note 299, at 601 (showing considerable success even within quite thin markets).

302 *See* KAY-YUT CHEN & CHARLES R. PLOTT, INFORMATION AGGREGATION MECHANISMS: CONCEPT, DESIGN, AND IMPLEMENTATION FOR A SALES FORECASTING PROBLEM 3 (Div. of the Humanities & Soc. Sci., Cal. Inst. of Tech., Social Science Working Paper No. 113, March 2002) (describing variation of this model employed by Hewlett-Packard), *available at* http://www.hss.caltech.edu/SSPapers/wp1131.pdf.

Imaged with Permission from N.Y.U. Law Review

3. The cost of an environmental regulation might be disputed and experts within government might be unsure about how to resolve the dispute. It would be possible to create an information market asking whether, by a specified date, the projected costs are above $400 million, above $500 million, or above $600 million.[303]

4. Much of the time, the benefits of environmental regulation are at least as controversial as the costs. Experts might disagree about whether a carcinogen causes harms at low levels, or even whether a substance is carcinogenic at all. An information market might be created to make predictions about the benefits of one or another course of action.[304]

5. Company executives might want to know which of their movies are most likely to be serious Oscar contenders. An Oscar nomination gives a large boost to ticket sales, and hence it is extremely valuable to be able to plan in advance. Existing markets might be enlisted to answer the relevant questions.[305]

6. Regulators are interested in trends involving air pollution, including increases or decreases in emissions over time and also in concentrations of pollutants in the ambient air. An information market might make projections about sulfur dioxide and particulate concentrations in New York City, Chicago, and Los Angeles in the next decade.

7. Officials in a political campaign, or in another nation, often want to know a particular candidate's likelihood of success at a particular moment in time. Instead of relying exclusively on polling data, they might consult information markets, assessing the likelihood of success for the candidate in question.

8. It is important both for government and for outside observers to know the size of federal budget deficits. Government projections are greatly disputed and some of them might well be

303 On the uses of information markets in cost-benefit analysis, see Abramowicz, *supra* note 299, at 997–1019.

304 *Id.* at 992–93. Of course it would be necessary to specify a source that would produce, at the relevant time, an authoritative judgment about benefits, that is, a judgment that could be deemed authoritative by all sides.

305 *See* Levmore, *supra* note 299, at 593 (describing how film studios have already begun utilizing information markets to structure film distribution).

Imaged with Permission from N.Y.U. Law Review

self-serving. Information markets might provide more reliable estimates.[306]

9. Regulators might be concerned about the likely risks of a new disease or of an old disease that seems to be growing in magnitude. To assess the risks, they might create an information market, designed to predict the numbers of deaths that will be attributed to, for example, mad cow disease over a specified period.

10. Federal and state agencies monitor a range of institutions to ensure that they are solvent.[307] One problem is that such agencies do not know whether insolvencies are likely to be many or few in a particular year; another is that the solvency of particular institutions can be difficult to predict in advance. Information markets could help with both problems.[308]

11. The national government might want to know the number of people who are likely to be infected by HIV in the United States or Africa by the year 2010; the answer to that question might be relevant to its policy judgments. An information market might be used to make predictions about the future progress of the disease.[309] Such markets might generally be used to make predictions about the likely effects of development projects, such as those involving vaccinations and mortality reductions.[310]

12. The government might seek to predict the likelihood and magnitude of damage from natural disasters, including tornadoes and earthquakes. Accurate information could greatly assist in advance planning. Information markets could easily be created to help in that task.

Some of these examples involve private behavior. Others involve the judgments of public institutions. Some might seem fanciful. Others involve predictions on which information markets are already flourishing. Let us turn, then, to actual practice.

---

306 *See* Abramowicz, *supra* note 299, at 990–92.

307 *See id.* at 987–90.

308 Most dramatically, they "might have led to earlier recognition of the savings and loans crisis in the 1980s." *Id.* at 988.

309 *See* ROBERT W. HAHN & PAUL C. TETLOCK, HARNESSING THE POWER OF INFORMATION: A NEW APPROACH TO ECONOMIC DEVELOPMENT 3–6 (AEI-Brookings Joint Ctr. For Regulatory Studies, Working Paper No. 04-21, 2004), *available at* http://www.aei-brookings.org/publications/abstract.php?pid=846.

310 *Id.*

Imaged with Permission from N.Y.U. Law Review

A. Practice and Evidence

1. An Abandoned Initiative

In many imaginable markets, people might make claims about facts, or predictions about the future, and they might stand to gain or lose from their predictions. In the summer of 2003, analysts at the Department of Defense built directly on this idea.[311] To predict important events in the world, including terrorist attacks, they sought to create a kind of market in which ordinary people actually could place bets. The proposed Policy Analysis Market potentially would have allowed people to invest in their predictions about such matters as the growth of the Egyptian economy, the death of Yassir Arafat, the military withdrawal of the United States from specified nations, and the likelihood of terrorist attacks in the United States. Investors would have won or lost money on the basis of the accuracy of their predictions.

Predictably, the Policy Analysis Market produced a storm of criticism. Ridiculed as "offensive" and "useless," the proposal was abandoned. Senator Tom Daschle called the market "the most irresponsible, outrageous and poorly thought-out of anything that I have heard the administration propose to date."[312] Senator Byron Dorgan argued that it is "morally bankrupt for a government agency to make a profitable game out of the deaths of American troops, heads of state, and nuclear missile attacks."[313] A private Policy Analysis Market, specializing in the Middle East, was promised in 2003, but it did not go forward.[314]

Amid the war on terrorism, why was the Department of Defense so interested in the Policy Analysis Market? The answer is simple: It wanted to have some assistance in predicting geopolitical events, including those that would endanger American interests, and it believed that a market would provide that help. It speculated that if a large number of people could be given an incentive to aggregate their private information, in the way that the Policy Analysis Market would do, government officials would learn a great deal. Apparently it believed that such a market would provide an important supplement

311 *See* Wolfers & Zitzewitz, *supra* note 299, at 107–08 (discussing this initiative).

312 Ronald Bailey, *Betting on Terror: Why Futures Markets in Terror and Assassinations Are a Good Idea*, REASONONLINE, July 30, 2003, *at* http://www.reason.com/rb/rb073003.shtml.

313 Byron L. Dorgan, Letter to the Editor, *The Pentagon's Ill-Conceived Market*, WASH. POST, Aug. 7, 2003, at A20.

314 For a replicate of the site, see Policy Analysis Market, *at* http://www.ratical.org/ratville/CAH/linkscopy/PAM (last visited Feb. 18, 2004).

Imaged with Permission from N.Y.U. Law Review

to deliberative processes both within government and without.[315] No one knows how the Department of Defense would have reacted to the projections of information markets; the most reasonable speculation is that those projections would have been used, not as the final word, but as providing valuable information about future events.

2. *Iowa Electronic Markets*

If this idea seems fanciful, consider the fact that since 1988, the University of Iowa has run the Iowa Electronic Markets (IEM), which allow people to bet on the outcome of presidential elections. Originally the IEM permitted people to trade only in the expected fraction of the popular vote to be obtained by presidential candidates.[316] Securities were offered that would pay $2.50 multiplied by the specified candidate's share of the vote. If, for example, George H.W. Bush received 50% of the vote, then the shareholder would receive $1.25. Shares could be bought and sold until the day before the election. Since their opening, the IEM have expanded from these modest roots. In the recent past, traders have been able to bet on the market capitalization that Google will achieve in its initial public offering, the price of Microsoft stock at a future date, and Federal Reserve monetary policy, in addition to betting on American elections.[317]

For presidential elections—still the most popular markets that IEM operates—traders have recently been permitted to choose from two types of markets.[318] In a "winner-take-all" market, traders win $1 for each "future" in the winning candidate that they own and nothing for shares of the losing candidate. In a "vote-share" market, traders in "candidate futures" win $1 multiplied by the proportion of the popular vote that the candidate received.[319] Thus, in a winner-take-all market, a "Dukakis future" was worth nothing after the election, while in a vote-share market, each Dukakis future paid $0.456. In a winner-take-all market, the market price reflects traders' perceptions

---

[315] *See* Wolfers & Zitzewitz, *supra* note 299, at 107 (discussing Defense Department proposal).

[316] *See* Joyce Berg et al., Accuracy and Forecast Standard Error of Prediction Markets 7–10 & nn.6–7 (July 2003) (unpublished manuscript), *available at* http://www.biz.uiowa.edu/iem/archive/forecasting.pdf.

[317] *See* Iowa Electronic Markets (operated by Henry B. Tippie Coll. of Bus., Univ. of Iowa), *at* http://www.biz.uiowa.edu/iem/markets (last visited Apr. 4, 2005).

[318] *See* Market Information: Pres04_WTA, Iowa Electronic Markets (operated by Henry B. Tippie Coll. of Bus., Univ. of Iowa), *at* http://128.255.244.60/WebEx/marketinfo_english.cfm?Market_ID=78 (last visited Feb. 18, 2005).

[319] *See* Market Information: Pres04_VS, Iowa Electronic Markets (operated by Henry B. Tippie Coll. of Bus., Univ. of Iowa), *at* http://128.255.244.60/WebEx/marketinfo_english.cfm?Market_ID=66 (last visited Feb. 18, 2005); *see also* Forsythe et al., *Wishes*, *supra* note 298, at 85 (discussing vote-share market and others).

Imaged with Permission from N.Y.U. Law Review

of the likelihood that each candidate will win the election. Perhaps more interestingly, observers can use the prices in a vote-share market much as they might use a poll. These prices are the market's estimate of each candidate's likely share of the vote when the election occurs. In each case, the market price reflects the aggregate information held by participants.

The IEM operate much like an ordinary stock market. To enter, each participant must purchase "unit portfolios" consisting of one future in each candidate for each dollar that the trader puts into the market.[320] Once she has bought enough of these "unit portfolios," she can unbundle the contracts and trade individual shares. All trading is fully computerized and traders must reach the markets through the Internet.[321] Unlike most stock exchanges, the IEM do not allow speculators to sell futures short. Nevertheless, as in a typical stock market, traders can issue bids and asks (limit orders) or accept outstanding offers (market orders). While most traders merely accept market orders rather than choosing their own prices, a small group of "marginal traders" trade frequently and post limit orders.[322] As we shall see, it is these traders who have the greatest effect on prices.

As a predictor, the Iowa Electronic Markets have produced extraordinarily accurate judgments. Before the 2004 elections, they did far better than professional polling organizations,[323] outperforming polls 451 out of 596 times.[324] In the week before the four elections from 1988 to 2000, the predictions in the Iowa market showed an average absolute error of just 1.5 percentage points, a significant improvement over the 2.1 percentage point error in the final Gallup polls.[325] In 2004, the Iowa market did even better, showing, on midnight of November 1, Bush with 50.45% of the vote, and Kerry with 49.55%—eerily close to the final numbers of 51.56% for Bush and 48.44% for Kerry.[326] This prediction was far better than the predictions that emerged from the more conventional indicators of likely results, including consumer confidence and job growth.[327]

---

320 Forsythe et al., *Wishes*, *supra* note 298, at 86.

321 *Id.*

322 *Id.* at 99–100.

323 *See* Wolfers & Zitzewitz, *supra* note 299, at 112.

324 *See* HAHN & TETLOCK, *supra* note 309, at 4.

325 *See* Wolfers & Zitzewitz, *supra* note 299, at 112.

326 *See* Erin Jordan, *Iowa Electronic Markets Yield Near-Accurate Result*, DES MOINES REG., Nov. 10, 2004, at 5B. Note that there was a great deal of volatility on election day, evidently produced by exit polls showing a likely victory by Kerry; I return to the issue of volatility below. *See id.*

327 Sue Kirchhoff, *Economic Predictors Don't Track Vote Results*, USA TODAY, Nov. 15, 2004, at 4B. Note, however, that "Polly," an aggregate predictor that includes the Iowa Electronic Markets, performed even better. *See supra* notes 72–73 and accompanying text.

Imaged with Permission from N.Y.U. Law Review

The IEM have proved accurate not only on election eve but also in long forecasting horizons.[328] Such markets are hardly limited to the United States. In other nations, universities are operating similar markets; examples include the University of British Columbia Election Stock Market, involving Canada,[329] and Vienna University of Technology, operating the Austrian Electronic Market.[330] Although the relevant districts are quite small, Australian bookmakers have shown a high degree of accuracy in predicting district-level races.[331] In fact, InTrade, a political market based in Dublin, accurately predicted not only the 2004 victory for President Bush but also the particular outcomes in the battleground states.[332]

3. *Other Information Markets: Hollywood, Weather, and Beyond*

Outside of the political context, consider the Hollywood Stock Exchange, in which people predict Oscar nominees and winners (as well as opening weekend box office successes). For the Hollywood Stock Exchange, the level of accuracy has been impressive, especially in view of the fact that the traders use virtual rather than real money. "HSX offers good predictions of a film's gross receipts before release and, relatively speaking, even better predictions after opening weekend—when a large number of traders have some information in the form of (or at least the possibility of) observing the finished film on screen, along with audience reactions."[333] As a result, "studios have begun relying on these estimates to structure the distribution of their films."[334] The market has proven successful in predicting award winners and box office returns.[335] Perhaps the most impressive achievement to date is its uncanny accuracy in predicting Oscar win-

328 *See* Berg et al., *supra* note 316, at 11–13.

329 Sauder School of Business, University of British Columbia, Election Stock Market, *at* http://esm.ubc.ca/ (last visited Apr. 6, 2005).

330 Technische Universität Wein, Austrian Political Stock Markets/Austrian Electronics Markets Homepage, *at* http://www.imw.tuwien.ac.at/apsm/ (last updated Mar. 12, 2003).

331 Justin Wolfers & Andrew Leigh, *Three Tools for Forecasting Federal Elections: Lessons from 2001*, 37 AUSTL. J. POL. SCI. 223, 234–40 (2002).

332 *See* John Tierney, *Now That the Dust Has Settled*, N.Y. TIMES, Nov. 7, 2004, § 1 (National) at 34.

333 Levmore, *supra* note 299, at 593.

334 *Id.*

335 Erica A. Klarreich, *Best Guess: Economists Explore Betting Markets as Prediction Tools*, 164 SCI. NEWS 251, 251 (2003), *available at* http://www.sciencenews.org/articles/2003 1018/bob9.asp; David M. Pennock et al., *The Real Power of Artificial Markets*, 291 SCIENCE 987, 987 (2001) (letter to the editor) (showing accuracy of Hollywood Stock Exhange).

Imaged with Permission from N.Y.U. Law Review

ners in 2005, with correct judgments in all eight of the categories for which trading was allowed.[336]

The success of information markets is seen in many other areas as well. For example, the futures market for oranges does a better job predicting weather in Florida than the National Weather Service.[337] A large prediction market, producing a typical event turnover in the hundreds of millions of dollars, focuses on the likelihood that economic data released later in the week will show specific values;[338] the market performs at least as well as the consensus forecasts of a survey of about fifty professional forecasters.[339] Companies have started to use internal prediction markets to answer relevant questions, including likely sales in specific periods.[340] The level of accuracy here is also high—far better, in all probability, than what would emerge from statistical means or deliberation, where excessive optimism can cause serious problems.

For example, Hewlett Packard (HP) and the California Institute of Technology initiated a project to study experimental markets as an information aggregation mechanism involving product sales.[341] The experimenters selected individuals who worked in different parts of HP's business operation. Because of its small size, the market was a very "thin" one, meaning that there were few participants and that the market was far less liquid than the much "thicker" Iowa Electronic Markets.[342] Participants were chosen with the thought that each could contribute information from his department in buying and selling the relevant futures, which were tied to sales and bonuses for executives

336 *See* Press Release, Hollywood Stock Exchange, Traders Hit a Perfect 100% of Oscar Winners (Feb. 28, 2005), *at* http://www.hsx.com/about/press/050229.htm. An intriguing puzzle: Less specialized markets, trading for real money and with substantial numbers of traders, proved less accurate than the Hollywood Stock Exchange. For an overview, see Stephen Phelan, *How to Win On Oscar*, SUNDAY HERALD, Feb. 27, 2005, at 12, which notes that according to the leading information markets, The Aviator would win for best picture, Clive Owen would win for best supporting actor, and Martin Scorsese would win for best director—all incorrect predictions. The puzzle lies in the fact that a smaller market, not involving real money, produced more accuracy than a larger one with significant stakes. One might expect that the real traders—on, for example, tradesports.com—would, at worst, see the trends on hsx.com, and invest accordingly. The solution to the puzzle probably lies in two facts: expertise and information costs. The traders on hsx.com are specialists, and the traders on the other markets, almost all of whom are not American, are generalists. If those generalists knew of the success of hsx.com, they would undoubtedly bet accordingly. I predict that in the future they will (but I'm not betting on it).

337 Richard Roll, *Orange Juice and Weather*, 74 AM. ECON. REV. 861, 871 (1984).

338 *See* Wolfers & Zitzewitz, *supra* note 299, at 113–14.

339 *Id.*

340 *See* Charles R. Plott, *Markets as Information Gathering Tools*, 67 S. ECON. J. 1, 12–13 (2000).

341 *See* CHEN & PLOTT, *supra* note 302, at 3.

342 *Id.* at 5–10.

Imaged with Permission from N.Y.U. Law Review

(which, in turn, are closely tied to profits).[343] The markets were organized so that securities existed for intervals of sales. For example, one security would pay off if sales were between one and ten printers; another would pay off if sales were between ten and twenty.[344] In most of the experiments, the possible range of sales was divided into ten intervals of equal size. On the basis of the prices of each security, the experimenters could guess how many units HP would sell that month. Information markets were expected to have large potential advantages over internal projections. Those involved in sales have an incentive to understate projected outcomes, so as to ensure that they do not fall short of expectations; this bias, or a competing bias in favor of excessive optimism, might well be reduced through market incentives.[345]

The results showed that the markets' predictions were a considerable improvement over HP's official forecasts. In no fewer than six of the eight markets for which official forecasts were available, the market prediction was significantly closer to the actual outcome than the official forecast[346]—and this was in spite of anecdotal evidence that the markets' activities were included as inputs in generating the official forecast.[347]

In fact, information markets are springing up all over the Internet, allowing people to make bets on the likely outcomes of sports, entertainment, finance, and political events.[348] We can find actual or proposed prediction markets about any number of questions: Will gas prices reach $3 per gallon? Will cellular life be found on Mars? Will Osama Bin Laden be captured by a certain date? Will small pox return to the United States? Will there be a sequel to Master and Commander? Will the Federal Communications Commission be abolished? These and other questions are being asked on information markets. Consider the following list:

- Hollywood Stock Exchange at http://www.hsx.com (site allowing users to bet on box office success of actors and movies).
- University of British Columbia Election Stock Market at http://esm.ubc.ca (market in which investors bet on outcomes of local and federal elections in Canada).

---

343 *Id.*

344 *Id.* at 7.

345 *Id.* at 3–4.

346 *Id.* at 12–13.

347 *Id.* at 5.

348 In many jurisdictions in the United States, however, these markets are forbidden as gambling; I do not explore the legal issues here.

Imaged with Permission from N.Y.U. Law Review

- Iowa Electronic Markets at http://www.biz.uiowa.edu/iem/ (market in which contract payoffs depend on economic and political events).
- Foresight Exchange at http://www.ideosphere.com/fx/ (entertainment site allowing users to bet on the likely outcome of future events using "funny money").
- Tradesports at http://www.tradesports.com (online trading exchange focused on the outcome of sporting events).
- News Futures at http://us.newsfutures.com/home/home.html (provider of prediction markets that delivers forecasts on issues for corporations).
- Probability Sports at http://www.probabilitysports.com (online sports betting site).
- Economic Derivatives at http://www.economicderivatives.com (online markets for economic derivatives).
- Wahlstreet at http://www.wahlstreet.de (German political futures market).

4. Aggregating Information Through Markets

All in all, prediction markets have been spectacularly successful in terms of the aggregate accuracy of the resulting "prices." Why is this? Note that such markets do not rely on the median or average judgment of a randomly selected group of people. They are genuine markets. Those who participate are self-selected. They must believe that they have relevant information; it is costly for them to "vote," and they probably will not do so unless they think that they have something to gain.[349] In addition, votes are not weighted equally. If people want to invest a few dollars, they are permitted to do so, but they can invest a great deal more if they are confident of their answer.[350] Intensity of belief is captured in prices.

There is a further point. People are permitted to buy and sell shares on a continuing basis. "Unlike polls . . . each trader in the

[349] Note that some markets involve virtual rather than real money. Newsfutures, for example, uses virtual currency that can be redeemed for monthly prizes (such as appliances); Foresight Exchange and the Hollywood Stock Exchange use "virtual currency," so that people do not earn real money, but instead attempt to enhance their reputation and their self-image. Note in this regard that Foresight Exchange lists publicly the "top ten investors by score." *See* Foresight Exchange Prediction Market, *at* http://www.ideosphere.com/fx/ (last visited Apr. 5, 2005).

[350] Some markets, however, impose limits on permissible investments; the IEM is an example, with a ceiling of $500. Frequently Asked Questions, Iowa Electronic Markets (operated by Henry B. Tippie Coll. of Bus., Univ. of Iowa), *at* http://www.biz.uiowa.edu/iem/faq.html#Questions (last visited Apr. 5, 2005). Note also that some markets, like the Hollywood Stock Exchange, do not involve real dollars. It is noteworthy that successful predictions are found even in such markets. Klarreich, *supra* note 335, at 251.

Imaged with Permission from N.Y.U. Law Review

market sees the net effect of the belief of all other traders, and the time series of changes in those beliefs . . . ."[351] As a a result, the market is "more than a static, one-time prediction but rather a dynamic system that can respond instantaneously to the arrival of new information."[352] Moreover, a correct answer is rewarded and an incorrect one is punished. Hence investors have a strong incentive to be right. In these circumstances, accurate answers can emerge even if only a small percentage of participants have good information. In the Iowa Electronic Markets, for example, it turns out that 85% of the traders do not seem to be particularly wise.[353] They hold onto their shares for a long period and then simply accept someone else's prices. The predictions of the market are driven by the other 15%—frequent traders who post their offers rather than accepting those made by other people. To work well, prediction markets do not require accurate judgments by anything like the majority of participants.[354] In this sense, information markets are very different from the ordinary judgments of deliberating groups. The resulting prices do not amplify or perpetuate cognitive errors; on the contrary, they correct them, because shrewd traders are able to invest in a way that corrects for even widespread errors.[355]

Of course information markets involve a measure of deliberation. Many individual investors deliberate with others before they invest. In some such markets, investors undoubtedly act as "teams," pooling resources after deliberating together about what to do. The point is that ultimate decisions come not from asking group members to come up with a mutually agreeable conclusion, but by reference to the price signal, which will have aggregated a great deal of diverse information. It is for this reason that information markets outperform deliberative processes.

---

351 Berg et al., *supra* note 316, at 9.

352 *Id.*

353 *See* Klarreich, *supra* note 335, at 252.

354 The same is, of course, true of ordinary markets. For a good overview, see ANDREI SHLEIFER, INEFFICIENT MARKETS: AN INTRODUCTION TO BEHAVIORAL FINANCE (2000).

355 As noted below, this is not inevitable. We could easily imagine a market in which cognitive problems are reflected in prices; indeed, this appears to happen with ordinary stock markets. *See generally* ROBERT J. SHILLER, IRRATIONAL EXUBERANCE (2000) (discussing relationship between cognitive errors and stock prices). In information markets, it is entirely possible to imagine booms or crashes produced by cognitive errors in combination with social influences. My point is not that this is impossible, but that the track record of information markets, at least thus far, is exceptionally good.

Imaged with Permission from N.Y.U. Law Review

5. *Building on Existing Practices*

How might institutions take advantage of information markets? It is possible to imagine both internal and public varieties. An internal market would be limited to people within the relevant organization. As we have seen, Hewlett Packard has used such a market to predict sales,[356] and the Department of Defense proposed an internal Policy Analysis Market as part of its abandoned initiative on geopolitical events. An external market would permit public investment by people outside of the institution for which predictions are being made. In either case, the outcome of the market might well be more accurate than the outcome of deliberation, in which errors might arise and be propagated or even amplified as a result of discussion.[357] An organization might rely on an internal market if it seeks to keep the results private or if it believes that an aggregation of information held within the organization will be sufficiently accurate. One risk of an internal market is that it might be too "thin," simply because most institutions will have few investors;[358] another is that members of the organization might suffer from a systematic bias. Alternatively, an institution might create a public market, available to all, believing that through this route it will obtain more accurate results. In either case, an organization might use an information market instead of group deliberation, or at the very least as an input into such deliberation.

B. Failed Predictions? Of Manipulation, Bias, and Bubbles

In what circumstances might information markets fail? To answer this question, ordinary stock markets are the place to start.[359] A great deal of recent attention has been paid to the possibility that individual traders are manipulable and also subject to identifiable biases in a way that leads them to blunder.[360] There is also a risk of

---

356 *See generally* CHEN & PLOTT, *supra* note 302.

357 For companies, optimistic bias is an obvious risk—one that information markets should reduce. *See* Daniel Kahneman & Dan Lovallo, *Timid Choices & Bold Forecasts: A Cognitive Perspective on Risk Taking*, 39 MGMT. SCI. (Jan. 1993) *reprinted in* CHOICES, VALUES & FRAMES 393, 393, 409–10 (Daniel Kahneman & Amos Tversky eds., 2000) (discussing decision makers' tendency to base forecasts of future outcomes on scenarios of success and to ignore past statistics). For an application to group decisions, see JOHNSON, *supra* note 25.

358 Note, however, that Hewlett Packard produced good predictions even in a thin market. CHEN & PLOTT, *supra* note 302, at 5, 12.

359 A general overview arguing that markets will not necessarily attain efficiency is SCHLEIFER, *supra* note 354.

360 *See, e.g.*, *id.*; HERSH SHEFRIN, BEYOND GREED & FEAR: UNDERSTANDING BEHAVIORAL FINANCE AND THE PSYCHOLOGY OF INVESTING (2000) (exploring markets' susceptibility to cognitive errors); SHILLER, *supra* note 355 (discussing cognitive errors and their

Imaged with Permission from N.Y.U. Law Review

"prediction bubbles," leading people to inaccurate judgments about future events.

1. *Manipulation*

A primary concern is that information markets, no less than ordinary ones, can be susceptible to manipulation by powerful speculators.[361] One attempt to manipulate an information market occurred during the 2000 presidential election. A group of speculators attempted to manipulate the Iowa Electronic Markets by buying large volumes of futures in presidential candidate Patrick Buchanan. The value of Buchanan shares did increase dramatically, but they fell almost immediately when "well-informed traders . . . seized the opportunity to profit off the manipulative traders . . . ."[362] Hence the Iowa market remained stable despite this attempted manipulation. Perhaps other, more plausible efforts at manipulation would succeed; but none have thus far.

2. *Biases*

Another concern is that some of the cognitive biases that afflict individuals will manifest themselves in prediction markets. Just as in group deliberation, investors in a market might be subject to predictable heuristics and biases.[363] The results here are unequivocal: They are subject to such effects. For example, psychologists have found that people overestimate the likelihood that their preferred candidate will win an election—a form of optimistic bias.[364] At a certain point in the 1980 campaign, 87% of Jimmy Carter's supporters believed that he would win, while 80% of Ronald Reagan's supporters believed that their candidate would win.[365] Obviously, at least one side had overestimated its candidate's probability of victory at the relevant time.

In the market context, similar biases can be found. Certain gamblers in New York are especially likely to bet on the New York Yankees;[366] and IEM traders show the same bias. In 1988, for

---

effects on market prices); ADVANCES IN BEHAVIORAL FINANCE (Richard H. Thaler ed., 1993) (investigating effects of how investors actually behave).

361 For a good discussion, see Abramowicz, *supra* note 299, at 972–76.

362 Klarreich, *supra* note 335, at 251, 253.

363 For an overview see SHILLER, *supra* note 355.

364 For an overview of optimistic bias, see Christine Jolls, *Behavioral Economics Analysis of Redistributive Legal Rules*, 51 VAND. L. REV. 1653, 1658–63 (1998).

365 Donald Granberg & Edward Brent, *When Prophesy Bends: The Preference-Expectation Link in U.S. Presidential Elections, 1952–1980*, 45 J. SOC. & PERSONALITY PSYCHOL. 477, 479 (1983).

366 *See* Wolfers & Zitzewitz, *supra* note 299, at 118 (citing Koleman S. Strumpf, Manipulating the Iowa Political Stock Market (2004) (unpublished mansuscript)).

Imaged with Permission from N.Y.U. Law Review

example, Dukakis supporters were more likely to hold futures in the Massachusetts governor's ill-fated presidential bid than were supporters of George H.W. Bush.[367] More strikingly still, Dukakis supporters were more likely to view the candidates' debates as helpful to the Democratic candidate and accordingly bought significant additional futures in his campaign after each debate.[368] Bush supporters showed precisely the same pattern. Traders thus exhibited the "assimilation-contrast" effect.[369] People usually assimilate new information in a way that confirms their view of the world, and those who invest in information markets show the same bias. In general, traders show a tendency to buy and sell in a way that fits with their party identification.[370]

Nonetheless, the Iowa Electronic Markets were more accurate than polls in predicting the outcome of the 1988 presidential election. Even three weeks before the election, the market provided an almost-perfect guess about the candidates' shares of the vote.[371] How is such accuracy possible when many traders showed identifiable biases? The answer may lie in the "marginal trader" hypothesis, about the behavior of a small group of "marginal traders" who were far less susceptible to these biases. According to this hypothesis, a small group of active traders who are far less susceptible to the relevant biases have a disproportionately large effect on aggregate market behavior. In trading election futures, these traders did not show the same biases as their fellow traders and earned significant profits at the expense of their quasi-rational colleagues.[372] Thus, the biased behavior of most traders did not affect the market price because the marginal traders were prepared to take advantage of their blunders. If marginal traders are active and able to profit from the bounded rationality of other participants, then there will be no effect on the aggregate market price.[373]

Another bias that might be expected to affect information markets is the "favorite-longshot" bias often seen in horse races. In horse racing, heavy favorites tend to give higher returns than other horses in

---

367 Forsythe et al., *Wishes*, *supra* note 298, at 94.

368 *Id.* at 94–95.

369 *See* MUZAFER SHERIF & CARL I. HOVLAND, SOCIAL JUDGMENT: ASSIMILATION AND CONTRAST EFFECTS IN COMMUNICATION AND ATTITUDE CHANGE 188 (1961) (discussing manner in which individuals filter information to conform to their pre-existing positions).

370 Forsythe et al., *Wishes*, *supra* note 298, at 94.

371 Berg et al., *supra* note 316, at 42.

372 Forsythe et al., *Wishes*, *supra* note 298, at 99–100. The term "quasi-rational" comes from RICHARD H. THALER, QUASI-RATIONAL ECONOMICS xxi (1991).

373 See the discussion in SCHLEIFER, *supra* note 354, at 24, 51–52.

Imaged with Permission from N.Y.U. Law Review

the field, while longshots tend to offer lower than expected returns.[374] Hence near certainties are undervalued while low probabilities are overvalued. If the point generalizes, prediction markets might not be accurate with respect to highly improbable events. The market should be expected to overestimate the likelihood that such events will come to fruition; for example, Patrick Buchanan futures would be expected to be (and might well have been) overpriced even before the attempted manipulation of the market. By contrast, an information market might underestimate the probability of events that are highly likely to occur. But with respect to existing prediction markets, there is only slight evidence of systematic errors in this vein.[375]

3. *Bubbles and More*

"Prediction bubbles" are also easy to imagine, with investors moving in a certain direction with the belief that many other investors are doing the same.[376] A temporary upsurge in investment in the nomination of Hillary Rodham Clinton as 2004 Democratic nominee might well have been a small bubble, with some investors thinking, not that she would in fact be the nominee, but that others would invest in that judgment, thus inflating the value of the investment. Crashes are possible as well. In any case informational influences certainly can lead individuals to make foolish investments in any market, including prediction markets.[377] As information markets develop, significant individual errors should be expected, and undoubtedly they will produce some errors in the price signal. In the 2004 presidential election, news of exit polls produced a great deal of volatility in election markets, with a dramatic election-day swing in the direction of Senator Kerry at the expense of President Bush. "Suddenly, Kerry's stock in the Winner Take All market shot up to 70 cents and Bush stock was in the cellar."[378] Large-scale errors are possible when apparently relevant news leads numerous investors to buy or sell; indeed, this particular shift may well have been a cascade, with inves-

---

374 *See* Richard H. Thaler & William T. Ziemba, *Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries*, 2 J. ECON. PERSP. 161, 163 (1988) (exploring favorite-longshot bias); *see also* Charles F. Manski, Interpreting the Predictions of Prediction Markets (Feb. 2004) (unpublished manuscript, on file with the *New York University Law Review*) (summarizing horse-race data findings), *available at* http://faculty.econ.nwu.edu/faculty/manski/prediction_markets.pdf.

375 The most important evidence can be found on Tradesports's predictions, where highly unlikely outcomes were overpriced in a number of domains. *See* Wolfers & Zitzewitz, *supra* note 299, at 117.

376 *See id.* at 118–19 (exploring speculative bubbles in context of information markets).

377 *Cf.* SHILLER, *supra* note 355 (discussing cognitive biases in trading context).

378 *See* Jordan, *supra* note 326.

Imaged with Permission from N.Y.U. Law Review

tors responding to one another's judgments, even though they were based on false information. The erroneous figures were able to last only for a few hours, however, after which the numbers returned to their previous state of considerable accuracy.

In particular contexts, the imaginable problems take a different form. Consider the problem of "terrorism futures." It would be extremely valuable to aggregate privately held information about the risk and location of any attack. But do likely investors actually possess helpful information? Thomas Rietz, a director of the Iowa Electronic Markets, argued that terrorism and world events were fundamentally different from other contexts in which markets have successfully predicted future events.[379] When betting on presidential elections, people can use ordinary information sources, along with their network of friends, family, and co-workers, to form an opinion; but for most investors, there are no such sources of information for terrorist activity. Another skeptic worried that the market would allow the wealthy to "hedge" against the possibility of terrorist activity, while ordinary Americans would remain vulnerable to this threat.[380] In this view, "terrorism futures" could operate as an insurance market that would not serve its purpose of providing information. In any event, government use of the resulting information could be self-defeating, at least if the information were made public. Terrorists would know the anticipated time and location of attacks, and also know that the government was aware of this—which would make it most unlikely that the prediction would turn out to be accurate. Where the event's occurrence is endogenous to the outcome of the information market, there is reason for skepticism about its likely performance, certainly if relevant actors have much to lose if the market turns out to be correct.[381]

But many policy issues, including those potentially involved in the now-defunct Policy Analysis Market, did not have this feature. Consider, for example, the question whether the Egyptian economy is likely to grow in the next year, or whether a Palestinian state will be created by the end of 2006. Perhaps many investors will lack a great deal of information on such questions, but it is most unlikely that the market prediction will turn out to be self-defeating. Of course the

379 Celeste Biever & Damian Carrington, *Pentagon Cancels Futures Market on Terror*, NEWSCIENTIST.COM, July 30, 2003, *at* http://www.newscientist.com/news/news.jsp?id=NS99994007.

380 Joseph E. Stiglitz, *Terrorism: There's No Futures in It*, L.A. TIMES, July 31, 2003 at B17.

381 *See* RICHARD A. POSNER, CATASTROPHE: RISK AND RESPONSE 175–76 (2004) (doubting usefulness of information markets in context of risk of terrorism).

Imaged with Permission from N.Y.U. Law Review

Policy Analysis Market itself raises many questions and doubts. And the growing body of work in behavioral finance gives reason for caution about excessive faith in information markets;[382] if ordinary stocks can be overvalued and undervalued, the same is undoubtedly true for predictions as well. The only point is that in many domains, information markets are extremely promising, and likely to outperform both statistical means and the products of group deliberation.

C. *Feasibility, Markets, and Deliberation Again*

Information markets face one pervasive problem: feasibility. A jury, for example, could not enlist such markets to decide on questions of guilt or innocence. Among other things, there is no objective way to test whether the jury, or individual jurors, ended up with the right answer (and if there were, the jury might be dispensable). Moreover, it is not easy to see how information markets could be used by judges. Of course factual questions are often relevant in court, but such markets could not easily be used to verify one or another answer. More generally, information markets might suffer from a legitimacy deficit, at least at the present time, where they remain unfamiliar. Recall that deliberation increases confidence and decreases variance; in many contexts, reliance on information markets might well breed confusion and distrust.[383]

There is another problem. When the relevant groups are small, effective markets may be impossible to create, simply because of the absence of sufficient numbers of investors.[384] A certain number is necessary to ensure that information markets have enough information to aggregate. On the other hand, administrative agencies might well enlist such markets,[385] and ambitious efforts are underway to examine how government might use them to answer an array of disputed questions.[386] At a minimum, such markets should be used, where feasible, as an adjunct to deliberative processes. Of course officials would not be bound by those predictions. They might reasonably believe that investors are wrong. But if the outcomes of information

---

382 *See* SCHLEIFER, *supra* note 354; SHILLER, *supra* note 355.

383 Recall the reaction to the Policy Analysis Market, outlined *supra* in the text accompanying notes 312–16.

384 Note, however, that "thin" markets have proved remarkably accurate, *see* Levmore, *supra* note 299, at 601–03 (discussing successes of thin markets), and that some small groups might encourage outsider investors.

385 *See* Abramowicz, *supra* note 299, at 992–93 (arguing for much greater use of information markets by administrative agencies).

386 *See* HAHN & TETLOCK, *supra* note 309, at 8 (arguing for multiple policy uses of information markets).

Imaged with Permission from N.Y.U. Law Review

markets prove reliable over time, officials should accept them unless they have grounds to believe that they are inaccurate.[387]

As Michael Abramowicz has suggested, governments might use information markets to help make projections about insolvency, budget deficits, and the costs and benefits of proposed regulations.[388] In each of these cases, deliberative processes might have a "reality check" in the form of the predictions of information markets. It would be possible to go much further. Officials might take into account the market's prediction of the anticipated damage from a natural disaster, the number of annual deaths from an actual or anticipated disease (such as mad cow disease or AIDS), the number of American casualties from a war effort, the existence of demonstrable harms from global warming by (say) 2010,[389] the likelihood of scarcity of natural resources,[390] decreases in tropical forests in the world, demonstrable deterrent effects from capital punishment or other severe punishments,[391] increases or decreases in emissions of specified air pollutants, increases or decreases in concentrations of air pollution in the ambient air, and much more. In all these cases, private or public institutions might create markets to provide information on

---

387 One complication here is that the market's prediction would be affected by the likelihood of government's response to that very prediction. If investors know, for example, that government is likely to respond if the market predicts a high number of fatalities from a natural disaster, then they should predict a lower number of fatalities, because government will by hypothesis be taking protective steps. This problem might be handled by conditional markets. For example, the question might be: "How many deaths will come from earthquakes in the United States in a specified year if government does not take new steps to prevent those deaths?"

388 Abramowicz, *supra* note 299, at 982–93.

389 For this example, and for others, it would of course be necessary to identify some agreed-upon source for an answer to the predictive question. We might specify, for example, that the judgment of the General Accounting Office is authoritative, rather than leaving investors to debate which source is most reliable.

390 Consider here the famous bet between Julian Simon and Paul Ehrlich about the coming scarcity of natural resources. Simon had long predicted that natural resources were essentially inexhaustible, whereas Ehrlich predicted that natural resources were running out. The two bet, in 1980, on the price of five metals to be selected by Ehrlich: If, by 1990, the price of the metals had risen (suggesting scarcity), Ehrlich would win. Ehrlich chose copper, chrome, nickel, tin, and tungsten. Ehrlich lost the bet. By 1990, the price of each of the five metals had fallen. For an overview, see Overpopulation.com, FAQ: People, Julian Simon's Bet with Paul Ehrlich, *at* http://www.overpopulation.com/faq/People/julian_simon.html (last visited Jan. 28, 2005).

391 As I have noted, it might be difficult, in some of these cases, to identify a source that would be deemed sufficiently objective. A public source—say, the General Accounting Office—might be specified as the relevant authority. For a relevant discussion on the difficulty in selecting data criteria, see Hashem Dezhbakhsh et al., *Does Capital Punishment Have a Deterrent Effect? New Evidence from Postmoratorium Panel Data*, 5 AM. L. & ECON. REV. 344, 355 (2004).

Imaged with Permission from N.Y.U. Law Review

crucial questions; and public institutions might take that information into account in making judgments about policy.

The broadest point is that, even when information markets are not feasible, an understanding of their virtues helps illuminate the virtues and vices of deliberation—and helps show how to obtain more of the former and less of the latter. Such markets overcome the collective action problem from which deliberating groups suffer; they also give people a strong incentive to say what they know and to back their most well-grounded convictions with money. An understanding of these points helps to show how to increase the likelihood that deliberation will not fall prey to the problems I have emphasized here.

V

NORMATIVE QUESTIONS AND GROUP JUDGMENTS

Deliberating groups are often asked to answer questions that are not purely factual. Issues involving morality, politics, and law require assessment of normative issues. Should cost-benefit analysis be the foundation of regulatory decisions? Should the minimum wage be increased? Should capital punishment be permitted? Can the President be impeached for lying under oath? Should *Roe v. Wade* be overruled? Should the Constitution be interpreted to require states to reconsider same-sex marriages? When, if ever, is theft morally acceptable?

When people answer such questions, informational influences and social pressures are likely to play a major role. One study demonstrates group polarization with respect to outrage: When individuals are outraged about corporate misconduct, juries become systematically more outraged than their median member.[392] And in fact group discussion often produces polarization on normative issues,[393] suggesting the presence of hidden profiles.[394] It is on normative questions, above all, that groups end up at a more extreme point in line with their predeliberation tendencies.[395] I have noted that in many domains, federal judges are subject to group polarization, with both Democratic and Republican appointees showing a tendency to

[392] *See* David Schkade et al., *supra* note 209, at 1155 (showing severity shift within juries).

[393] *See* BROWN, *supra* note 22, at 220–26 (discussing group polarization); SUNSTEIN, *supra* note 21, at 120–24 (exploring sources of polarization).

[394] If groups are polarizing toward an extreme position, it may well be because certain information, not widely shared within the group, remains hidden.

[395] *See* BROWN, *supra* note 22, at 222–26.

Imaged with Permission from N.Y.U. Law Review

extremism when they are sitting with like-minded others.[396] We might therefore attempt to find ways to reduce the potential problems of deliberation when groups are exploring normative questions, not simply questions of fact.

It might be controversial to suggest that groups amplify individual errors in the normative domain, because in that domain, we might not be able to say with confidence that one or another view counts as an "error." Skeptics about morality and law, rejecting the view that moral and legal questions have correct answers, would insist that any shifts introduced by deliberation cannot be said to be right or wrong. But if they are correct, does deliberation have any point?[397] In any case, skepticism is extremely hard to defend for law or morality. We may bracket the debate over whether legal problems have uniquely correct answers in hard cases[398] while also agreeing that, on multiple and diverse views about legal reasoning, some conclusions are right and others are wrong.[399] And if deliberation is often likely to lead people to err on questions of fact, it will also lead participants in law to blunder on questions of law. Suppose, for example, that the question is whether a regulatory agency has violated the statute that it is charged with administering, or whether a particular voting scheme violates the Equal Protection Clause, or whether the impossibility doctrine relieves a contracting party of the duty to perform. In all of these cases, groups are likely to err if their deliberations are not structured in such a way as to overcome the risks of amplification of errors, hidden profiles, cascade effects, and group polarization.

In the moral domain, skepticism also runs into serious problems.[400] Without engaging the complex philosophical issues, we

---

396 *See* Sunstein, Schkade & Ellman, *supra* note 23, at 305 (finding polarization in voting by federal judges).

397 A possible answer would stress the legitimating functions of deliberation, *see supra* note 9 and accompanying text, but for deliberation to work, and even to legitimate, deliberators must believe that they are trying to make progress on a disputed question, not simply to legitimate it. An effort to justify deliberation purely on the ground that it is legitimating will tend to be self-defeating for the participants.

398 *See* RONALD DWORKIN, A MATTER OF PRINCIPLE 144–45 (1985) (arguing that legal disputes have right answers).

399 This proposition follows, for example, from views as diverse as those expressed by Ronald Dworkin, Antonin Scalia, and Cass Sunstein. *See* RONALD DWORKIN, LAW'S EMPIRE 255 (1986) (defending idea of "integrity," in which judges attempt to decide legal issues on the basis of broad principles); ANTONIN SCALIA, A MATTER OF INTERPRETATION 23–24 (1997) (defending textualism); CASS R. SUNSTEIN, LEGAL REASONING AND POLITICAL CONFLICT 35–44 (1996) (defending use of incompletely theorized agreements in law).

400 Perspectives on skepticism vary widely. *See generally* DAVID O. BRINK, MORAL REALISM AND THE FOUNDATIONS OF ETHICS (1989) (defending moral realism as foundation for ethics); GILBERT HARMON & JUDITH JARVIS THOMPSON, MORAL RELATIVISM AND MORAL OBJECTIVITY (1996) (critiquing arguments for moral skepticism); JOHN

Imaged with Permission from N.Y.U. Law Review

can simply note that many different views about the nature of morality acknowledge the possibility of individual error—and that if individual error does occur, group error will occur as well. As obvious examples, consider the persistence of slavery and racial segregation. As a less obvious example, consider the fact that people's answers to many questions depend on how those questions are framed. The framing of options affects judgments not only on factual questions but on moral ones as well, including, for example, the disputed issue of moral obligations to members of future generations.[401] As noted, groups do not show less susceptibility to framing effects than individuals;[402] it follows that groups will be vulnerable to framing for questions of morality and law as well as for questions of fact.

No information market could be helpful in answering normative questions, simply because there is no way to establish whether a particular investor was correct; for normative questions, predictions are not being made at all.[403] And for such questions, it might seem odd or perhaps even bizarre to rely on the judgments of statistical groups. To be sure, democratic processes might be seen as an effort to settle moral and political issues by seeking the mean view within the relevant population (views that are formed after deliberation, at least much of the time). But to say the least, it is controversial to claim that moral and political questions are best answered by simply finding the mean views of a population-wide sample. Is the morality of abortion, or capital punishment, properly settled by asking for the average view of a group of, say, one thousand people? Is a legal question to be resolved by taking the median view of a large set of people trained in the law?

Ordinarily moral and legal answers are found by reference to the reasons offered on behalf of competing positions, not by taking a poll. Suppose it is true that the average judgment is accurate when large groups are asked about the number of beans in a jar, the weight of a large animal, or the temperature in a room. Perhaps it follows that the average judgment will be accurate when large groups are asked about abortion, capital punishment, and preemptive war. But systematic biases or errors may infect the judgments of group members,

---

RAWLS, A THEORY OF JUSTICE 48–53 (1971) (discussing search for reflective equilibrium); BERNARD WILLIAMS, *Interlude: Relativism*, *in* MORALITY: AN INTRODUCTION TO ETHICS 20, 20–26 (1972) (condemning functionalist view of moral relativism).

401 *See* Cass R. Sunstein, *Moral Heuristics and Moral Framing*, 88 MINN. L. REV. 1556, 1591–94 (2004) (discussing effects of framing in context of rights of future generations).

402 *See* Kerr et al., *supra* note 64, at 688, 698–702.

403 It might be tempting to say that the moral views of posterity provide the relevant test, but then the bet would be on the moral views of posterity, not on what morality requires.

Imaged with Permission from N.Y.U. Law Review

making the Condorcet Jury Theorem inapplicable, and there is no simple way to test whether this is so.

Note, however, that empirical questions are often a central component of good answers to normative problems; many such problems cannot be resolved without knowing something about the facts. The analysis of mistakes by deliberating groups should apply in full force to the factual components of normative questions. Consider, for example, the suggestion that the minimum wage should be increased. If minimum wage increases would significantly decrease employment, surely that is relevant to the decision whether to support such increases; and it matters too whether minimum wage increases would benefit poor people or mostly people who are not poor.[404]

To be sure, these are empirical questions which experts are almost certainly far better able to answer accurately than deliberating groups of ordinary people. The point is only that many normative questions cannot sensibly be resolved without information about the actual effects of one or another answer. When this is so, an understanding of the hazards of deliberation, and how to minimize those hazards, can be used constructively by groups that are attempting to resolve normative questions. Of course consequences may not be the central part of some normative disputes. Some people believe, for example, that capital punishment is morally unacceptable even if it has a strong effect in deterring murders, and evaluative judgments of various kinds can separate people even if they agree on the facts.[405] But the more general point nonetheless holds: Sometimes a certain view of the facts can bring diverse people into line on normative issues, producing a single position despite disagreements on those issues. To this extent, the analysis here applies to normative questions as well. Group judgments on such questions will be distorted by the amplification of errors, hidden profiles, cascade effects, and polarization. It is important to take steps, of the kind that I have catalogued, to reduce those distortions.

What about for purely normative issues, lacking any factual component? Here the argument on behalf of group deliberation is not fundamentally different from what it is elsewhere.[406] Unless we are

---

404 For an analysis of empirical evidence on who benefits from minimum wage legislation, see Daniel Shaviro, *The Minimum Wage, the Earned Income Tax Credit, and Optimal Subsidy Policy*, 64 U. CHI. L. REV. 405, 433–39 (1997).

405 *See* Dan M. Kahan & Donald Braman, *More Statistics, Less Persuasion: A Cultural Theory of Gun-Risk Perceptions*, 151 U. PA. L. REV. 1291, 1308 (2003) (emphasizing role of culture in producing perceptions of risks).

406 *See* GUTMANN & THOMPSON, *supra* note 4, at 1–9 (discussing virtues of deliberative conception of democracy).

Imaged with Permission from N.Y.U. Law Review

relativists or skeptics, we will agree that one point of deliberation is to ensure that normative questions are correctly answered, that is, are answered by reference to good reasons, even if we disagree about what they are. And if this is so, then there is strong reason to be concerned, for normative questions no less than empirical ones, that group judgments will be impaired by the mechanisms traced here. The structural reforms have an equivalent role in the normative domain. We may therefore take the simple cases I have emphasized, in which deliberation leads to palpable and demonstrable errors, to provide clear evidence of deliberative pathologies that are likely to occur even when errors are neither palpable nor demonstrable. If a central goal is to ensure that normative questions—in law, politics, and morality—are answered well, then the prescriptions I have outlined deserve a place in numerous deliberating groups, including those not centrally concerned with facts at all.

A more general lesson follows. Many people have celebrated deliberative conceptions of democracy, largely with the thought that deliberative processes are likely to lead participants in better or more sensible directions. The underlying judgment is easy to understand. New perspectives can broaden old ones, ensuring that people see far more, as a result of deliberation, than they saw before. Everyone has had the experience, on questions of value as well as of fact, of learning from the claims offered by others. But if the argument here is correct, that experience can be misleading. Even under ideal conditions, emphasized by proponents of deliberation,[407] group members can be led to err, not despite deliberation but because of it. When members of a religious community end up entrenched in their belief in (say) traditional gender roles, or when members of a liberal group end up still more committed to a highly progressive income tax, social dynamics, leaving hidden profiles, may well be responsible. The fact that deliberation increases confidence, and reduces variance, will leave the strong but potentially misleading impression that deliberation has produced sense rather than nonsense.

Nothing I have said here is meant to suggest that deliberation should be eliminated, or that on normative questions, groups would do better to rely on some kind of statistical average. The proper response to the problems I have traced is to structure deliberative processes so as to ensure that relevant points of view will emerge as a result of those processes. With values as well as facts, it is crucial to prime critical thinking and to encourage group members to believe that cooperation, in the form of quiescence, can be greatly overrated.

---

[407] *See generally* HABERMAS, *supra* note 4; GUTMANN & THOMPSON, *supra* note 4.

Imaged with Permission from N.Y.U. Law Review

In the abstract, advice of this sort risks vacuousness. But I hope that I have said enough to show that such advice can be used as the basis for concrete institutional recommendations, designed to ensure that even like-minded groups contain mechanisms that promote a kind of internal system of checks and balances.

CONCLUSION

Groups often contain a great deal of information, and an important task is to elicit and use the information that members actually have. Deliberation is generally thought to be the best way of carrying out that task. But deliberative bodies are subject to serious problems. Much of the time, informational influences and social pressures lead members not to say what they know. As a result, groups tend to propagate, and even to amplify, cognitive errors. They also emphasize shared information at the expense of unshared information; hidden profiles are a result. Cascade effects and group polarization are also common. Those interested in deliberative democracy have yet to pay adequate attention to these problems, which are likely to lead deliberating groups to blunder even if participants behave in accordance with the conditions of the deliberative ideal. Here, then, is a context in which political philosophy does not engage with existing work, both theoretical and empirical, on individual and social cognition.

What can be done by way of response? At the very least, it should be possible to structure deliberation so as to increase the likelihood that relevant information will emerge. A norm in favor of critical thinking, and incentives to reward individuals for good decisions by groups, can overcome some of the relevant pressures. Leaders should take steps to encourage a wide range of views; to do this, leaders might be cautious about expressing their own views at the outset and should encourage reasons, rather than conclusions, before the views of group members start to harden. Institutions might ensure anonymity and private polling before deliberation; they might permit anonymous statements of final conclusions; they might create strong incentives, economic and otherwise, to encourage people to disclose what they know.

Information markets have significant advantages over deliberative processes, and in many contexts they might supplement or even replace those processes. Such markets tend to correct rather than to amplify individual errors, above all because they allow shrewd inves-

Imaged with Permission from N.Y.U. Law Review

tors to take advantage of the mistakes made by others.[408] Because information markets provide economic rewards for correct individual answers, they realign incentives in a way that promotes disclosure. As a result, they are often more accurate than the judgments of deliberating groups. To the extent feasible, many groups would often do well to enlist information markets in arriving at their judgments, above all because of the accuracy of the price signal.

My emphasis throughout has been on the aggregation of information and the risk that deliberating groups will err on instrumental questions and on issues of fact. But the same risks arise in the normative domain, where informational influences and social pressures also produce forms of self-silencing that are highly damaging to good deliberation. In that domain, as elsewhere, it makes no sense to celebrate deliberation in the abstract. Incentives make all the difference. Well-functioning groups take steps to ensure that on normative questions, as on factual ones, people feel free to disclose what they believe to be true.

---

[408] Note that this is an empirical claim, not a conceptual one. It is certainly possible for markets to propagate individual errors. *See generally* SHILLER, *supra* note 355 (discussing role of cognitive errors in stock market).

Imaged with Permission from N.Y.U. Law Review