

Essay

Deliberative Trouble? Why Groups Go to Extremes

Cass R. Sunstein[†]

The differences of opinion, and the jarring of parties in [the legislative] department of the government . . . often promote deliberation and circumspection, and serve to check excesses in the majority.

—Alexander Hamilton¹

In everyday life the exchange of opinion with others checks our partiality and widens our perspective; we are made to see things from their standpoint and the limits of our vision are brought home to us. . . . The benefits from discussion lie in the fact that even representative legislators are limited in knowledge and the ability to reason. No one of them knows everything the others know, or can make all the same inferences that they can draw in concert. Discussion is a way of combining information and enlarging the range of arguments.

—John Rawls²

[†] Karl N. Llewellyn Distinguished Service Professor of Jurisprudence, University of Chicago Law School and Department of Political Science. I am grateful above all to Daniel Kahneman and David Schkade for many helpful discussions. Thanks too to C. Edwin Baker, Robert Ellickson, Edward Glaeser, Robert Goodin, Christine Jolls, Dan Kahan, Timur Kuran, Lawrence Lessig, John Manning, Andrei Marmor, Wiktor Osiatynski, Eric Posner, Richard Posner, Jeffrey Rachlinski, and David Strauss for valuable comments. Participants in workshops at Cornell and the University of Chicago provided a great deal of help. I am also grateful to participants in a conference on deliberative democracy at the University of Texas at Austin, and in particular to my commentators Norman Bradburn and Robert Luskin. Brian Lehman and Brooke May provided excellent research assistance.

1. THE FEDERALIST NO. 70, at 426-27 (Alexander Hamilton) (Clinton Rossiter ed., 1961).

2. JOHN RAWLS, A THEORY OF JUSTICE 358-59 (1971).

Each person can share what he or she knows with the others, making the whole at least equal to the sum of the parts. Unfortunately, this is often not what happens

. . . As polarization gets underway, the group members become more reluctant to bring up items of information they have about the subject that might contradict the emerging group consensus. The result is a biased discussion in which the group has no opportunity to consider all the facts, because the members are not bringing them up.

. . . .

. . . Each item they contributed would thus reinforce the march toward group consensus rather than add complications and fuel debate.

—Patricia Wallace³

I. INTRODUCTION

Consider the following events:

- Affirmative action is under attack in the state of Texas. A number of professors at a particular branch of the University of Texas are inclined to be supportive of affirmative action; they meet to exchange views and, if necessary, to plan further action. What are these professors likely to think, and to do, after they talk?
- After a nationally publicized shooting at a high school, a group of people in the community, most of them tentatively in favor of greater gun control, come together to discuss the possibility of imposing new gun control measures. What, if anything, will happen to individual views as a result of this discussion?
- A jury is deciding on an appropriate punitive damage award in a case of recklessly negligent behavior by a large company; the behavior resulted in a serious injury to a small child. Before deliberating as a group, individual jurors have chosen appropriate awards, leading to an average of \$1.5 million and a median of \$1 million. As a statistical generalization, how will the jury's ultimate award tend to compare to these figures?

3. PATRICIA WALLACE, *THE PSYCHOLOGY OF THE INTERNET* 81-82 (1999).

- A group of women is concerned about what they consider to be a mounting “tyranny of feminism.” They believe that women should be able to make their own choices, but they also think that men and women are fundamentally different, and that their differences legitimately lead to different social roles. The group decides to meet every two weeks to focus on common concerns. Is it possible to say what its members are likely to think after a year?

Every society contains innumerable deliberating groups. Church groups, political parties, women’s organizations, juries, legislative bodies, regulatory commissions, multimember courts, faculties, student organizations, those participating in talk radio programs, Internet discussion groups, and others engage in deliberation. It is a simple social fact that sometimes people enter discussions with one view and leave with another, even on moral and political questions.⁴ Emphasizing this fact, many recent observers have embraced the traditional American aspiration to “deliberative democracy,” an ideal that is designed to combine popular responsiveness with a high degree of reflection and exchange among people with competing views.⁵ But for the most part, the resulting literature has not been empirically informed.⁶ It has not dealt much with the real-world consequences of deliberation, and with what generalizations hold in actual deliberative settings, with groups of different predispositions and compositions.

The standard view of deliberation is that of Hamilton and Rawls, as quoted above. Group discussion is likely to lead to better outcomes, if only because competing views are stated and exchanged. Aristotle spoke in similar terms, suggesting that when diverse people

all come together . . . they may surpass—collectively and as a body, although not individually—the quality of the few best . . . [W]hen

4. Sometimes it may seem that moral and political arguments are unlikely to have an effect on individual judgments; the evidence discussed here shows that arguments often induce considerable movement.

5. *E.g.*, DELIBERATIVE DEMOCRACY (Jon Elster ed., 1998); AMY GUTMANN & DENNIS THOMPSON, DEMOCRACY AND DISAGREEMENT 128-64 (1996); JÜRGEN HABERMAS, BETWEEN FACTS AND NORMS 287-328 (William Rehg trans., Polity Press 1996) (1992); CASS R. SUNSTEIN, THE PARTIAL CONSTITUTION 133-45 (1993).

6. Exceptions include James D. Fearon, *Deliberation as Discussion*, in DELIBERATIVE DEMOCRACY, *supra* note 5, at 44; and Susan C. Stokes, *Pathologies of Deliberation*, in DELIBERATIVE DEMOCRACY, *supra* note 5, at 123. Of special interest are James Fishkin’s continuing experiments with the “deliberative opinion poll,” in which groups of diverse people are asked to deliberate on public issues. See James S. Fishkin & Robert C. Luskin, *Bringing Deliberation to the Democratic Dialogue*, in THE POLL WITH A HUMAN FACE 3 (Maxwell McCombs & Amy Reynolds eds., 1999). Fishkin’s groups do not polarize, at least not systematically; this result is undoubtedly a product of the distinctive setting, in which materials are presented on each issue, with corresponding claims of fact and value. For a discussion of Fishkin, see *infra* text accompanying notes 193-203.

there are many [who contribute to the process of deliberation], each has his share of goodness and practical wisdom . . . [S]ome appreciate one part, some another, and all together appreciate all.⁷

But an important empirical question is whether and under what circumstances it is really true that “some appreciate one part, some another, and all together appreciate all.”

My principal purpose in this Essay is to investigate a striking but largely neglected⁸ statistical regularity—that of group polarization—and to relate this phenomenon to underlying questions about the role of deliberation in the “public sphere”⁹ of a heterogeneous democracy. In brief, group polarization means that members of a deliberating group predictably move toward a more extreme point in the direction indicated by the members’ predeliberation tendencies.¹⁰ “[L]ike polarized molecules, group members become even more aligned in the direction they were

7. ARISTOTLE, THE POLITICS 108-09 (Ernest Barker trans., Oxford Univ. Press 1995).

8. Though the topic has been much studied within psychology, e.g., ROGER BROWN, SOCIAL PSYCHOLOGY: THE SECOND EDITION 200-48 (1986), I have been unable to find sustained discussions in the relevant literature in economics, sociology, philosophy, law, or political science, and there appears to be no treatment of the implications of group polarization for social behavior or the theory of democracy. Within the economics literature, an exception is Timothy Cason & Vai-Lam Mui, *A Laboratory Study of Group Polarisation in the Team Dictator Game*, 107 ECON. J. 1465 (1997), described by the authors as “the first study that attempts to incorporate this psychological literature into economics.” *Id.* at 1466. For one of the rare treatments within sociology, see Noah E. Friedkin, *Choice Shift and Group Polarization*, 64 AM. SOC. REV. 856 (1999). Group polarization is also explored in Kenneth L. Bettenhausen, *Five Years of Groups Research: What We Have Learned and What Needs To Be Addressed*, 17 J. MGMT. 345, 356-59 (1991).

There have been some valuable discussions of group polarization in the context of jury deliberations. A 1976 jury simulation study, BROWN, *supra*, at 226-29, tests for, and finds, group polarization, *id.* at 228; and a number of other mock jury studies, not intended to test for group polarization, collect evidence that it occurred in “every instance where the report of data makes it possible to check.” *Id.* at 229; see also David Schkade et al., *Deliberating About Dollars: The Severity Shift*, 100 COLUM. L. REV. 1139 (2000).

9. The term is associated with Jürgen Habermas. HABERMAS, *supra* note 5, at 360-87; JÜRGEN HABERMAS, THE STRUCTURAL TRANSFORMATION OF THE PUBLIC SPHERE (Thomas Burger trans., Mass. Inst. of Tech. Press 1989) (1962).

10. Note that this statement has two different implications. First, a deliberating group, asked to make a group decision, will shift toward a more extreme point in the direction indicated by the median predeliberation judgment. Second, the tendency of individuals who compose a deliberating group, if polled anonymously after discussion, will be to shift toward a more extreme point in the direction indicated by the median predeliberation judgment. Frequently these two phenomena are collapsed in the empirical literature, and I will not always distinguish between them here. But for some purposes it is important to distinguish them, and hence some work refers to the movement of groups as “choice shifts” and the movement of individuals as “group polarization.” E.g., Johannes A. Zuber et al., *Choice Shift and Group Polarization: An Analysis of the Status of Arguments and Social Decision Schemes*, 62 J. PERSONALITY & SOC. PSYCHOL. 50, 50, 59 (1992). I discuss this distinction in more detail below. *Infra* text accompanying notes 64-65.

Note also that in the experimental work, both extremism and tendencies are measured not by reference to anything external, nor to a normative standard, but by reference to the particular scale that is brought before the individuals who compose the group. Thus, for example, people might be asked, on a scale of -5 to 5, how strongly they agree or disagree with a particular statement.

already tending.”¹¹ Thus, for example, members of the first deliberating group are likely to become more firmly committed to affirmative action; the second group will probably end up favoring gun control quite enthusiastically; the punitive damages jury will likely come up with an award higher than the median, perhaps higher than the average as well, and very possibly as high as or higher than that of the highest predeliberation award of any individual member; and the group of women concerned about feminism is likely to become very conservative indeed on gender issues. Notably, groups consisting of individuals with extremist tendencies are more likely to shift, and likely to shift more; the same is true for groups with some kind of salient shared identity (like Republicans, Democrats, and lawyers, but unlike jurors and experimental subjects).¹² When like-minded people are participating in “iterated polarization games”—when they meet regularly, without sustained exposure to competing views—extreme movements are all the more likely.

Two principal mechanisms underlie group polarization. The first points to social influences on behavior and in particular to people’s desire to maintain their reputation and their self-conception. The second emphasizes the limited “argument pools” within any group, and the directions in which those limited pools lead group members. An understanding of the two mechanisms provides many insights into deliberating bodies. Such an understanding illuminates a great deal, for example, about likely processes within multimember courts, juries, political parties, and legislatures—not to mention ethnic groups, extremist organizations, criminal conspiracies, student associations, faculties, institutions engaged in feuds or “turf battles,” workplaces, and families. At the same time, these mechanisms raise serious questions about deliberation from the normative point of view.¹³ If deliberation predictably pushes groups toward a more extreme point in the direction of their original tendency, whatever that tendency may be, is there any reason to think that deliberation is producing improvements?

One of my largest purposes is to cast light on *enclave deliberation* as simultaneously a potential danger to social stability, a source of social fragmentation, and a safeguard against social injustice and unreasonableness.¹⁴ Group polarization helps explain an old point, with

11. JOHN C. TURNER ET AL., REDISCOVERING THE SOCIAL GROUP 142 (1987).

12. Dominic Abrams et al., *Knowing What To Think by Knowing Who You Are*, 29 BRIT. J. SOC. PSYCHOL. 97, 112 (1990).

13. I am speaking here of real-world deliberation, not of deliberation accompanied by preconditions of the sort that have been identified by those thinking of it in ideal terms. See, e.g., HABERMAS, *supra* note 5, at 99-131.

14. On the last point, see the discussion of African-American newspapers in RONALD JACOBS, RACE, MEDIA, AND THE CRISIS OF CIVIL SOCIETY: FROM WATTS TO RODNEY KING 140-51 (2000).

clear constitutional resonances, to the effect that social homogeneity can be quite damaging to good deliberation.¹⁵ When people are hearing echoes of their own voices, the consequence may be far more than support and reinforcement. An understanding of group polarization thus illuminates social practices designed to reduce the risks of deliberation limited to like-minded people. Consider the ban on single-party domination of independent regulatory agencies, the requirement of legislative bicameralism, and debates, within the United States and internationally, about the value of proportional or group representation. Group polarization is naturally taken as a reason for skepticism about enclave deliberation and for seeking to ensure deliberation among a wide group of diverse people.

But there is a point more supportive of enclave deliberation: Participants in heterogeneous groups tend to give least weight to the views of low-status members¹⁶—in some times and places, women, African Americans, less-educated people. Hence enclave deliberation might be the only way to ensure that those views are developed and eventually heard. Without a place for enclave deliberation, citizens in the broader public sphere may move in certain directions, even extreme directions, precisely because opposing voices are not heard at all. The ambivalent lesson is that deliberating enclaves can be breeding grounds both for the development of unjustly suppressed views and for unjustified extremism, indeed fanaticism.

This Essay is organized as follows. Part II discusses social influences on individual judgments, with particular reference to the phenomenon of social cascades. When a number of people have acted or spoken, observers who lack much private information are highly likely to follow their lead, a phenomenon that can drive participants in politics in unexpected and sometimes extreme directions. Part III offers a basic account of group polarization, with an elaboration of the underlying mechanisms of social influence and persuasive arguments. Parts IV, V, and VI deal with the implications of group polarization for democracy and law. Part IV is descriptive; it explores the likely consequences of group polarization for a number of issues at the intersection of law and democratic theory, including feuds, ethnic strife, juries, commissions, multimember courts, and deliberation via the Internet. Part V is normative; it draws the various strands together in an effort to show how group polarization can create deliberative trouble. It also traces the implications of the phenomenon for structuring deliberative institutions and links group polarization to central aspects of constitutional design. Part VI focuses on the role of deliberation inside and outside enclaves of like-minded people, emphasizing the need to

15. For the constitutional resonances, see *infra* notes 164-165, 169-171.

16. Caryn Christensen & Ann S. Abbott, *Team Medical Decision Making*, in *DECISION MAKING IN HEALTH CARE* 267, 273-76 (Gretchen Chapman & Frank Sonnenberg eds., 2000).

ensure that members of deliberating enclaves are not walled off from other points of view. Part VII is a brief conclusion.

II. SOCIAL INFLUENCES AND CASCADES

A. *In General*

A great deal of attention has recently been devoted to the topic of social influences on individual behavior.¹⁷ Because many of these influences are analogous to what happens in group polarization, and because they have a bearing on democratic deliberation as well, it is worthwhile to offer some brief notations here.

Social influences can lead groups to go quite rapidly in identifiable directions, often as a result of “cascade” effects involving either the spread of information (whether true or false) or growing reputational pressure. Sometimes cascade effects are highly localized, and lead members of particular groups, quite rationally,¹⁸ to believe or to do something that members of other groups, also quite rationally, find to be baseless or worse. Local cascades can ensure that different groups end up with very different, but equally entrenched, views about the same issues and events.¹⁹

People frequently think and do what they think and do because of what they think relevant others think and do.²⁰ Thus, for example, employees are more likely to file suit if members of the same work group have also done so;²¹ littering and nonlittering behavior appear to be contagious;²² a good way to increase the incidence of tax compliance is to inform people of high

17. E.g., Dan M. Kahan, *Social Influence, Social Meaning, and Deterrence*, 83 VA. L. REV. 349 (1997). For extended overviews, see ELLIOT ARONSON, *THE SOCIAL ANIMAL* (7th ed. 1995); and LEE ROSS & RICHARD E. NISBETT, *THE PERSON AND THE SITUATION* (1991). Group polarization is a surprising omission from both of these lengthy and highly illuminating treatments.

18. On the rationality of participation in a cascade, see David Hirschleifer, *The Blind Leading the Blind: Social Influence, Fads, and Informational Cascades*, in *THE NEW ECONOMICS OF HUMAN BEHAVIOR* 188, 189 (Mariano Tommasi & Kathryn Ierulli eds., 1995).

19. For an entertaining and candid discussion, see the outline of attitudes of Microsoft employees, by a Microsoft employee, in Michael Kinsley, *The View from the Cafeteria: Microsoft Employees Don't Recognize Themselves in the Government's Suit*, *TIME*, Apr. 10, 2000, at 152, 2000 WL 17632849; the discussion can reasonably be taken to show both cascade effects and group polarization.

20. ROSS & NISBETT, *supra* note 17, at 28-57; Dan M. Kahan, *Gentle Nudges vs. Hard Shoves: Solving the Sticky Norms Problem*, 67 U. CHI. L. REV. 607 (2000); Kahan, *supra* note 17.

21. Harold H. Gardner et al., *Workers' Compensation and Family and Medical Leave Act Claim Contagion*, 20 J. RISK & UNCERTAINTY 89, 101-10 (2000).

22. Robert B. Cialdini et al., *A Focus Theory of Normative Conduct: Recycling the Concept of Norms To Reduce Littering in Public Places*, 58 J. PERSONALITY & SOC. PSYCHOL. 1015, 1017 (1990).

levels of voluntary tax compliance;²³ television networks tend to follow one another;²⁴ and students are less likely to engage in binge drinking if they think that most of their fellow students do not engage in binge drinking, so much so that disclosure of low numbers of binge drinkers is one of the few successful methods of reducing binge drinking on college campuses.²⁵

Social influences affect behavior via two different mechanisms.²⁶ The first is informational. What other people do, or say, carries an *informational externality*;²⁷ if many other people support a particular candidate, or refuse to use drugs, or carry guns, observers, and particularly observers within a common group, are given a signal about what it makes sense to do. The second mechanism is reputational, as group members impose sanctions on perceived deviants, and would-be deviants anticipate the sanctions in advance.²⁸ Even when people do not believe that what other people do provides information about what actually should be done, they may think that the actions of others provide information about what other people *think* should be done. Thus each person's expressive actions come with a *reputational externality*. People care about their reputations and they have an incentive to do what (they think) other group members think they should do. For example, reputational considerations may lead people to obey or not to obey the law, urge a certain view in group discussions, buy certain cars,²⁹ drive while drunk, help others, or talk about political issues in a certain way. A concern for reputation exerts a ubiquitous influence on behavior, including that of participants in democratic debate, who often shift their public statements in accordance with reputational incentives.³⁰

23. STEPHEN COLEMAN, MINN. DEP'T OF REVENUE, THE MINNESOTA INCOME TAX COMPLIANCE EXPERIMENT STATE TAX RESULTS 1, 5-6, 18-19 (1996), <http://www.state.mn.us/ebranch/mdor/reports/compliance/pdf>.

24. Robert E. Kennedy, Strategy Fads and Competitive Convergence: An Empirical Test for Herd Behavior in Prime-Time Television Programming (Aug. 1999) (unpublished manuscript, on file with *The Yale Law Journal*).

25. Timur Kuran & Cass R. Sunstein, *Availability Cascades and Risk Regulation*, 51 STAN. L. REV. 683, 767 (1999); H. Wesley Perkins, *College Student Misperceptions of Alcohol and Other Drug Norms Among Peers: Exploring Causes, Consequences, and Implications for Prevention Programs* 177, 194, in U.S. DEP'T OF EDUC., DESIGNING ALCOHOL AND OTHER DRUG PREVENTION PROGRAMS IN HIGHER EDUCATION: BRINGING THEORY INTO PRACTICE (1997), available at <http://www.edc.org/hec/pubs/theorybook/perkins.pdf>. A good outline of contagion effects can be found in Gardner et al., *supra* note 21, at 91-93.

26. E.g., ARONSON, *supra* note 17, at 22; ROSS & NISBETT, *supra* note 17, at 44-45.

27. Andrew Caplin & John Leahy, *Miracle on Sixth Avenue: Information Externalities and Search*, 108 ECON. J. 60, 61 (1998).

28. GEORGE A. AKERLOF, *A Theory of Social Custom, of Which Unemployment May Be One Consequence*, in AN ECONOMIC THEORIST'S BOOK OF TALES 69 (1984).

29. See ROBERT H. FRANK, LUXURY FEVER 8-10, 122-45 (1999) (connecting luxury purchases to competition for better relative position). This is only part of Frank's account; he also emphasizes the "frame of reference" created by most people's decisions. *Id.* at 131-32.

30. Kuran & Sunstein, *supra* note 25, at 691-703, 715-35 (discussing the relationship between certain cascades and democratic politics). See generally TIMUR KURAN, PRIVATE TRUTH, PUBLIC LIES 3-21 (1995) (discussing preference falsification because of public pressures).

B. *Some Classic Experiments*

In the most vivid experiments involving group influences, conducted by Solomon Asch, individuals were apparently willing to abandon the direct evidence of their own senses.³¹ In the relevant experiments, a certain line was placed on a large white card. The task of the subjects was to match that line by choosing, as identical to it in length, one of three other lines that had been placed on a separate large white card. One of the lines on the second white card was in fact identical in length to the line to be matched to it; the other two were substantially different, with the differential varying from an inch and three-quarters to three-quarters of an inch. The subject in the original experiments was one of eight people asked to engage in the matching. Unbeknownst to the subject, the other people apparently being tested were actually there as Asch's confederates, serving as part of the experiments.

In the first two rounds of the experiment, everyone agreed about the right answer; this seemed to be an extremely dull experiment. But the third round introduced "an unexpected disturbance."³² Other group members made what was obviously, to the subject and to any reasonable person, a clear error. They matched the line at issue to one that was obviously longer or shorter. In these circumstances the subject had the choice of maintaining his independent judgment or instead yielding to the crowd. A large number of people ended up yielding. In ordinary circumstances, subjects erred less than 1% of the time, but in rounds in which group pressure supported the incorrect answer, subjects erred 36.8% of the time.³³ Indeed, in a series of twelve questions, no fewer than 70% of subjects went along with the group, and defied the evidence of their own senses, at least once.

Susceptibility to group influence was hardly uniform. Some people agreed with the group almost all of the time, whereas others were entirely independent in their judgments. Significantly, a modest variation in the experimental conditions made all the difference. The existence of at least one compatriot, or voice of sanity, dramatically reduced both conformity and error. When just one other person made an accurate match, errors were reduced by three-quarters, even if there was a strong majority the other way.³⁴ By contrast, varying the size of the group unanimously making the erroneous decision mattered only up to a number of three; increases from that point had little effect. Opposition from one person did not increase subjects' errors at all; opposition from two people increased error to 13.6%;

31. For an overview, see Solomon E. Asch, *Opinions and Social Pressure*, in READINGS ABOUT THE SOCIAL ANIMAL 13 (Elliot Aronson ed., 1995).

32. *Id.* at 15.

33. *Id.* at 16.

34. *Id.* at 18.

and opposition from three people increased error to 31.8%, not substantially different from the level that emerged from further increases in group size.

More recent studies have identified an important feature of social influence, directly bearing on group behavior in democracies. Much depends on the subject's perceived relationship to the experimenters' confederates and in particular on whether the subject considers himself part of the same group in which those confederates fall. Thus conformity—and error—is dramatically *increased* in public statements when the subject perceives himself as part of a reasonably discrete group that includes the experimenter's confederates (all psychology majors, for example).³⁵ By contrast, conformity is dramatically *decreased*, and hence error is also dramatically decreased, in public statements when the subject perceives himself as in a different group from the experimenter's confederates (all ancient history majors, for example).³⁶ Notably, private opinions, expressed anonymously afterwards, were about the same whether or not the subject perceived himself as a member of the same group as others in the experiment. Apparently, public statements of agreement with a majority view are particularly likely to be both inaccurate and insincere when relevant speakers identify themselves as members of the same group as the majority.³⁷

Both informational and reputational considerations appear to lead people toward the relevant errors. In Asch's own studies several people said, in private interviews, that their own opinions must have been wrong.³⁸ On the other hand, these statements may have been an effort to avoid the dissonance that would come from confessing that the statement was false but made only for reputational reasons. Notably, experimenters find greatly reduced error when the subject is asked to give a purely private answer.³⁹ And in the study described in the immediately preceding paragraph, people who thought that they were members of the same group as the

35. Abrams et al., *supra* note 12, at 106-08. See also the discussion of the “downside” of social ties among group members in Brooke Harrington, *Cohesion, Conflict and Group Demography* (1999) (unpublished manuscript, on file with *The Yale Law Journal*), which shows that when social ties are in place, dissent may be suppressed and decisions may be worse as a result.

36. Abrams et al., *supra* note 12, at 106-08.

37. The least conformity, and the greatest accuracy, were found when people who thought of themselves as being in a different group were speaking publicly. At the same time, the largest number of conforming, inaccurate responses came when people thought of themselves as being in the same group and were speaking publicly—even though the number of inaccurate *private* responses in that experimental condition was not notably higher than in other conditions. *Id.* at 108.

38. Asch, *supra* note 31, at 17.

39. ARONSON, *supra* note 17, at 23-24.

experimenter's confederates gave far more accurate answers, and far less conforming answers, when they were speaking privately.⁴⁰

Asch concluded that his results raised the possibility that "the social process is polluted" by the "dominance of conformity. . . . That we have found the tendency to conformity in our society so strong that reasonably intelligent and well-meaning young people are willing to call white black is a matter of concern."⁴¹ Asch's experiments did not involve deliberation, for people were not exchanging reasons; indeed, one might expect that reasoning on the part of Asch's confederates would have lessened the amount of conformity and error. What reasons could have been given for incorrect matches? But the existence of substantial numbers of mistakes as a result of mere exposure to the incorrect conclusions of others raises questions about whether and when deliberation within groups will lead people in the right directions.

C. *Social and Law-Related Cascades*

A great deal of recent work on social influence has raised the possibility of informational and reputational "cascades";⁴² this work has obvious relevance to law and politics.⁴³ (Consider issues involving race and sex equality, global warming, capital punishment, AIDS, the filing of lawsuits, or presidential candidates.) Indeed, Asch's work demonstrates considerable individual susceptibility to cascade effects. What is striking about such effects is their epidemic-like nature, or the quality of apparent contagion. Group polarization is sometimes, but not always, a product of cascade effects. It is useful to understand the former against the background of the latter.

40. Abrams et al., *supra* note 12, at 108. By contrast, people who thought that they were members of a different group actually gave more accurate, nonconforming answers when speaking *publicly*, which creates an interesting puzzle: Why was there more accuracy in public than in private statements? The puzzle is solved by considering the likelihood that subjects could consider it an affirmative good to disagree with people from another group (even if they secretly suspected that those people might be right). In the real world, this effect may well be heightened when people are asked whether they agree with opponents or antagonists; they might well say "no" even when the answer is "yes," simply because agreement carries costs, either to reputation or to self-conception.

41. Asch, *supra* note 31, at 21.

42. E.g., Lisa R. Anderson & Charles A. Holt, *Information Cascades in the Laboratory*, 87 AM. ECON. REV. 847 (1997); Abhijit V. Banerjee, *A Simple Model of Herd Behavior*, 107 Q.J. ECON. 797 (1992); Sushil Bikhchandani et al., *Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades*, 12 J. ECON. PERSP. 151 (1998); Andrew F. Daughety & Jennifer F. Reinganum, *Stampede to Judgment: Persuasive Influence and Herding Behavior by Courts*, 1 AM. L. & ECON. REV. 158, 159-65 (1999).

43. See Kuran & Sunstein, *supra* note 25, at 763-67.

1. *Informational Cascades*

When individuals lack a great deal of private information (and sometimes even when they have such information), they rely on information provided by the statements or actions of others. If *A* is unaware whether abandoned toxic waste dumps are in fact hazardous, he may be moved in the direction of fear if *B* seems to think that fear is justified. If *A* and *B* both believe that fear is justified, *C* may end up thinking so too, at least if she lacks reliable independent information to the contrary. If *A*, *B*, and *C* believe that abandoned toxic waste dumps are hazardous, *D* will need a good deal of confidence to reject their shared conclusion.

People typically have different thresholds for choosing to believe or do something new or different. As those with low thresholds come to a certain belief or action, people with somewhat higher thresholds join them, possibly to a point where a critical mass is reached, making groups, possibly even nations, “tip.”⁴⁴ The result of this process can be to produce snowball or cascade effects, as small or even large groups of people end up believing something—even if that something is false—simply because other people seem to believe that it is true. There is a great deal of experimental evidence of informational cascades, which are easy to induce in the laboratory;⁴⁵ real-world phenomena also seem to have a great deal to do with cascade effects.⁴⁶ Consider, for example, smoking, participating in protests, striking, rioting, buying stocks,⁴⁷ choosing what to put on television,⁴⁸ even leaving bad dinner parties.⁴⁹

Though the cascades phenomenon has been discussed largely in connection with factual judgments, the same processes should be at work for political, legal, and moral questions. Suppose, for example, that *A* believes that affirmative action is wrong and even unconstitutional, that *B* is otherwise in equipoise but shifts upon hearing what *A* believes, and that *C*

44. Mark Granovetter, *Threshold Models of Collective Behavior*, 83 AM. J. SOC. 1420, 1441-42 (1978). For a recent popular treatment, see MALCOLM GLADWELL, *THE TIPPING POINT* 5-22 (2000).

45. See generally Anderson & Holt, *supra* note 42.

46. See Sushil Bikhchandani et al., *A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades*, 100 J. POL. ECON. 992 (1992); Kuran & Sunstein, *supra* note 25, at 715-35.

47. ROBERT J. SHILLER, *IRRATIONAL EXUBERANCE* 151-67 (2000).

48. Kennedy, *supra* note 24.

49. Several of these examples are discussed in Kuran & Sunstein, *supra* note 25, at 725-35; and Granovetter, *supra* note 44, at 1422-24. With respect to AIDS, unfortunate and literally deadly cascade effects can be found in South Africa. President Mbeke's widely publicized doubts about a connection between AIDS and HIV came from learning about the views of “denialists” as a result of surfing the Internet. Thomas H. Maugh II, *AIDS Researchers Meet at Ground Zero*, L.A. TIMES, July 10, 2000, at A1, 2000 WL 2258962. In addition, many myths have spread, suggesting, for example, that there is a “miracle cure” and that organic food is a shield or a cure. See Donald G. McNeil Jr., *South Africa: Aw C'mon, You Don't Really Believe Those AIDS Myths*, AFRICA NEWS, June 11, 1999, 1999 WL 19530508.

is unwilling to persist in his modest approval of affirmative action when *A* and *B* disapprove of it. It would be a very confident *D* who would reject the judgments of three (apparently) firmly committed others. Sometimes people are not entirely sure whether capital punishment should be imposed, whether the Constitution protects the right to have an abortion, whether it is wrong to litter or to smoke. Many people, lacking firm convictions of their own, end up believing what relevant others seem to believe.

Judges are also vulnerable to cascade effects.⁵⁰ The same process is sometimes at work for the choice of political candidates, as a fad develops in favor of one or another—a cascade “up” or “down,” with sensational or ruinous consequences. One can easily imagine cascade effects in the direction of certain judgments about the appropriate course of government regulation, the minimum wage, or even constitutional law; indeed, such effects seem to have been at work in the legal culture in the 1960s (with mounting enthusiasm for the Warren Court) and the 1980s (with mounting skepticism about that Court). It is even possible to imagine cascade effects with respect to questions of constitutional method (for example, textualism, originalism, democracy-reinforcing judicial review). Note that a precondition for an informational cascade is a lack of much private information on the part of many or most people. If people have a good deal of private information, or are confident about their own judgments, they are unlikely to be susceptible to the signals sent by the actions of others.⁵¹

2. *Reputational Cascades*

There can be reputational pressures and reputational cascades as well,⁵² in which people speak out, or remain silent, or even engage in certain expressive activity, partly in order to preserve their reputations, at the price of failing to say what they really think. Suppose, for example, that *A* believes that hazardous waste dumps pose a serious environmental problem; suppose too that *B* is skeptical. *B* might keep quiet, or (like some of Asch’s subjects) even agree with *A*, simply in order to preserve *A*’s good opinion. *C* might see that *A* believes that hazardous waste dumps pose a serious problem and that *B* seems to agree with *A*; *C* might therefore voice agreement even though privately she is skeptical or ambivalent.

It is easy to see how this kind of thing might happen in political life—with, for example, politicians expressing their commitment to capital punishment (even if they are privately skeptical) or their belief in God

50. See Daughety & Reinganum, *supra* note 42, at 167-82 (discussing the possibility of herd behavior by courts).

51. This point has an echo in findings on group polarization, as those with entrenched views are less likely to “move” as a result of discussion. See *infra* text accompanying notes 98-99.

52. KURAN, *supra* note 30, at 4-21.

(even if they are privately agnostic). People typically have different thresholds for yielding to perceived reputational pressure. Some people follow perceived pressure only when it is very severe (for example, because a large number of people impose it), whereas others will follow when it is mild (for example, simply because some relevant others impose it). Here too the consequence can be cascade effects—large social movements in one direction or another—as increasing numbers of people yield to a pressure that they simultaneously impose, eventually reaching a critical mass. At that stage, a large number of people eventually appear to support a certain course of action simply because others (appear to) do so.

As in the context of the informational cascade, what is true for factual beliefs may be true as well for moral, legal, and political judgments. People might say, for example, that affirmative action violates the Constitution simply because of perceived reputational sanctions from saying the opposite; they might support or oppose the death penalty largely in order to avoid the forms of social opprobrium that might come, in the relevant community, from taking the opposing view. A precondition for reputational cascades is that in making the decision at issue, reputational considerations loom large. If people do not care about their reputations, or if reputation is a small component of the choice involved, the perceived intrinsic merits will be crucial, and cascades are unlikely to result.

Are social cascades good or bad? No general answer would make sense. Sometimes cascades are quite fragile, precisely because people's commitments are based on little private information; sometimes cascades are rooted in (and greatly fuel) blunders. Sometimes cascade effects will eliminate group or public torpor by generating concern about serious though previously ignored problems. But sometimes cascade effects will make people far more worried than they should be, or otherwise produce large-scale distortions in private judgments, public policy, and law. It is reasonable to speculate that both the antislavery and the Nazi movements had distinctive cascade-like features, as did the environmental movement in the United States and the fall of communism.⁵³ The serious risk with social cascades, both informational and reputational, is that they can lead to widespread errors, factual or otherwise. Cascades need not involve deliberation, but related problems infect processes of group deliberation, as I show below.

53. See Susanne Lohmann, *The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-91*, 47 *WORLD POL.* 42, 76 (1994).

III. HOW AND WHY GROUPS POLARIZE

How do small groups of like-minded people differ from large groups of heterogeneous people? What is likely to happen within isolated deliberating enclaves? How does all this bear on deliberative democracy? To answer these questions, and to understand the relationship between social processes within groups⁵⁴ and democratic theory, it is necessary to have an understanding of group polarization.

A. *The Basic Phenomenon*

Group polarization is among the most robust patterns found in deliberating bodies, and it has been found in many diverse tasks. Polarization “is said to occur when an initial tendency of individual group members toward a given direction is enhanced following group discussion.”⁵⁵ The result is that groups often make more extreme decisions than would the typical or average individual in the group (where “extreme” is defined solely internally, by reference to the group’s initial dispositions). There is a clear relationship between group polarization and cascade effects: Both have a great deal to do with informational and reputational influences. An important difference is that group polarization, unlike cascade effects, involves deliberation. In addition, polarization may or may not involve a cascade-like process; polarization can result simply from simultaneous independent decisions to move toward the group extreme.

Though standard, the term “group polarization” is somewhat misleading. It is not meant to suggest that group members will shift to two poles, nor does it refer to an increase in variance among groups, though this may be the ultimate result.⁵⁶ Instead the term refers to a predictable shift *within* a group discussing a case or problem. As the shift occurs, groups and group members move and coalesce, not toward the middle of antecedent dispositions, but toward a more extreme position in the direction indicated by those dispositions. The effect of deliberation is both to decrease variance among group members, as individual differences diminish, and also to

54. I do not deal here with the general and related question whether groups amplify or eliminate various cognitive and motivational biases in individual decisions. For a general overview, finding mixed results, see Norbert L. Kerr et al., *Bias in Judgment: Comparing Individuals and Groups*, 103 PSYCHOL. REV. 687 (1996).

55. Daniel J. Isenberg, *Group Polarization: A Critical Review and Meta-Analysis*, 50 J. PERSONALITY & SOC. PSYCHOL. 1141, 1141 (1986).

56. When different deliberating groups polarize in different directions, the consequence can be greater variance among groups, notwithstanding small initial differences. A group whose members are initially but tentatively disposed to reject some proposal might start out very close to a group whose members are initially but tentatively disposed to approve of that proposal. If the two groups have a number of intragroup discussions, but no intergroup discussions, they may end up very far apart.

produce convergence on a relatively more extreme point among predeliberation judgments.

The basic phenomenon has been found in many nations.⁵⁷ Consider some examples:

- After discussion, citizens of France become more critical of the United States and its intentions with respect to economic aid.⁵⁸
- A group of moderately profeminist women becomes more strongly profeminist after discussion.⁵⁹
- As a result of deliberation, whites predisposed to greater racial prejudice offer more negative responses to the question of whether white racism is responsible for conditions faced by African Americans in American cities.⁶⁰
- After discussion, whites predisposed not to show racial prejudice offer more positive responses to the same question.⁶¹

In all of these cases, deliberation moves people toward a more extreme point in line with their predeliberation positions.⁶²

B. *Risky Shifts and Cautious Shifts*

Group polarization was first found in a series of experiments involving risk-taking decisions. These experiments found what has become known as the “risky shift.”⁶³ In the original experiments, male graduate students of

57. BROWN, *supra* note 8, at 222-26. These include the United States, England, and France. *See also, e.g.*, Abrams et al., *supra* note 12, at 112 (New Zealand); Zuber et al., *supra* note 10 (Germany). Of course, it is possible that some cultures would show a greater or lesser tendency toward polarization; this would be an extremely interesting area for empirical study.

58. BROWN, *supra* note 8, at 224.

59. David G. Myers, *Discussion-Induced Attitude Polarization*, 28 HUM. REL. 699, 710-12 (1975).

60. BROWN, *supra* note 8, at 224 (discussing a study in David G. Myers & George D. Bishop, *The Enhancement of Dominant Attitudes in Group Discussion*, 20 J. PERSONALITY & SOC. PSYCHOL. 386 (1971)).

61. *Id.*

62. In laboratory studies, polarization occurs in terms of a specified issue and a specified scale. Thus any tendency toward extremism is scale-relative; it need not mean extremism as a normative matter, or across a full population. In the real world, political entrepreneurs, with self-interested or altruistic agendas, are in some sense aware of the importance of the scale at issue, and attempt to produce shifts along the scale that has been self-consciously made salient. The salient scales (“Do you approve of quotas?” as opposed to “Do you believe in remedial steps for traditionally disadvantaged groups?”) can be extremely important.

63. BROWN, *supra* note 8, at 248 (citing J.A.F. Stoner, *A Comparison of Individual and Group Decisions Including Risk* (1961) (unpublished M.S. thesis, Massachusetts Institute of Technology) (on file with the Massachusetts Institute of Technology Library)); James A.F. Stoner,

industrial management were asked a range of questions involving risk. For twelve of the thirteen groups, the group decisions showed a repeated pattern toward greater risk-taking. After discussion, for example, the unanimous outcome assessed the acceptable probability of financial soundness as consistently lower than the median judgment of group members prior to deliberation. In addition, there was a clear shift toward greater risk-taking in private opinions. This shift—the “risky” shift—was promptly duplicated in a number of diverse studies, some involving all men and some involving all women.

It is important to distinguish at this point between two aspects of these findings⁶⁴ usually not separated in the psychological literature and both relevant to law and policy. The first involves the movement of deliberating groups, when a group decision is necessary, toward the group’s extreme end; call this a *choice shift*. This means that if a group decision is required, the group will tend toward an extreme point, given the original distribution of individual views. Undoubtedly the group’s decision rule will matter here; a requirement of unanimity, for example, may well produce a shift toward the most extreme points, at least if those with the most extreme views are least tractable and most confident. The second involves the movement of (even private) individual judgments as a result of group influence; call this simply (as is standard) group polarization. This means that to the extent that the private judgments of individuals are moved by discussion, the movement will be toward a more extreme point in the direction set by the original distribution of views. It is possible to have one kind of movement without the other, though ordinarily the two accompany one another,⁶⁵ as with risky shifts.

Studies have shown that under certain conditions, it is possible, even easy, to induce a “cautious shift” as well. Indeed, certain problems reliably produce cautious shifts.⁶⁶ The principal examples involve the decision whether to marry and the decision whether to board a plane despite severe abdominal pain possibly requiring medical attention. In these cases, deliberating groups move toward caution, as do the members who compose them. Burglars, in fact, show cautious shifts in discussions, though when they work together, the tendency is toward greater risk-taking.⁶⁷

In Stoner’s original data, researchers noticed that the largest risky shifts could be found when group members “had a quite extreme risky initial

Risky and Cautious Shifts in Group Decisions, 4 J. EXPERIMENTAL SOC. PSYCHOL. 442 (1968). In this and the following paragraph, I draw from BROWN, *supra* note 8, at 200-06.

64. For an insistence on this distinction, see Zuber et al., *supra* note 10, at 50, 59.

65. *Id.* at 59 (“[I]t is necessary to differentiate between choice shift and group polarization . . . [T]he ongoing processes at these two levels may differ.”).

66. See BROWN, *supra* note 8, at 208-10, for an overview.

67. Paul F. Cromwell et al., *Group Effects on Decision-Making by Burglars*, 69 PSYCHOL. REP. 579, 586 (1991).

position,” in the sense that the predeliberation votes were weighted toward the risky end, whereas the items “that shifted a little or not at all started out near the middle of the scale.”⁶⁸ Thus the direction of the shift seemed to turn on the location of the original disposition, and the size of the shift depended on the extremeness of that original disposition. Similar results have been found in many contexts relevant to law and democracy, involving, for example, questions about economic aid, architecture, education, political leaders, race, and judgments of guilt or innocence.⁶⁹ Polarization has been found for questions of obscure fact as well as for evaluative questions, including political and legal issues⁷⁰ and even the attractiveness of people in slides.⁷¹

C. *Mechanisms*

What explains group polarization? It is tempting to think that conformity plays a large role, and as the Asch experiments suggest, individual judgments can be greatly influenced by the desire to conform. Perhaps conformity is sometimes at work, but the data make it clear that group polarization is not a matter of conformity; people do not shift to the mean of initial positions.⁷² The relevant movement goes to one or another side. Indeed, this is what defines, and what is most distinctive about, group polarization.

There have been two main explanations for group polarization. Both of these have been extensively investigated and supported.⁷³ The first explanation of group polarization—social comparison—begins with the claim that people want to be perceived favorably by other group members and also to perceive themselves favorably. Most people seek to take a position of a certain socially preferred sort. In the case of risk-taking, for example, they may want to be perceived (and to perceive themselves) as daring risk-takers, and their choice of position is partly a product of these

68. BROWN, *supra* note 8, at 211.

69. *Id.* at 224, 226-27.

70. A relatively recent treatment is Russell Spears et al., *De-Individuation and Group Polarization in Computer-Mediated Communication*, 29 BRIT. J. SOC. PSYCHOL. 121 (1990).

71. TURNER ET AL., *supra* note 11, at 153.

72. BROWN, *supra* note 8, at 207-08. Note, however, that one account of group polarization finds that the effect lies in conformity to the “prototypical group member,” defined as such by reference to a “meta-contrast principle: the less a person differs from in-group members and the more he or she differs from out-group members, the more representative is he or she of the in-group.” Craig McGarty et al., *Group Polarization as Conformity to the Prototypical Group Member*, 31 BRIT. J. SOC. PSYCHOL. 1, 3 (1992). This position raises many questions; it seems to have the strongest fit with the data in cases in which in-groups and out-groups can readily be understood as such by subjects. *See id.*

73. BROWN, *supra* note 8, at 210-29, and Isenberg, *supra* note 55, review this literature; see also TURNER ET AL., *supra* note 11, at 142-70, for an overview and an attempt to generate a new synthesis.

desires.⁷⁴ No one can know what the desirable position would be until the positions of others are revealed.⁷⁵ People move their judgments in order to preserve their image to others and their image to themselves. Evidence confirms that mere exposure to the positions of others, without discussion, induces a substantial risky shift (though this shift is only about half as large as the shift produced by discussion).⁷⁶ This effect helps explain a shift toward caution (the “cautious shift”) as well.⁷⁷

The second explanation is based on the commonsense intuition that any individual’s position on an issue is partly a function of which arguments presented within the group seem convincing. The choice therefore moves in the direction of the most persuasive position defended by the group, taken as a whole. Because a group whose members are already inclined in a certain direction will have a disproportionate number of arguments going in that same direction, the result of discussion will be to move people further in the direction of their initial inclinations. The key is the existence of a limited argument pool, one that is skewed (speaking purely descriptively) in a particular direction.⁷⁸ Hence there will be a shift in the direction of the

74. For a quite vivid demonstration of a related process in the enactment of the Clean Air Act—one that does not, however, identify the mechanisms discussed here—see E. Donald Elliott et al., *Toward a Theory of Statutory Evolution: The Federalization of Environmental Law*, 1 J.L. ECON. & ORG. 313 (1985).

75. Brown notes:

Once the real location of the mean was known, should it not be the case, granting that everyone wanted to see himself as reasonably audacious, that those who were really below the mean would be motivated to adopt riskier positions and so change the mean and produce the risky shift?

BROWN, *supra* note 8, at 214.

76. Allan I. Teger & Dean G. Pruitt, *Components of Group Risk Taking*, 3 J. EXPERIMENTAL SOC. PSYCHOL. 189 (1967).

77. One might point here to both one-upmanship and the removal of pluralistic ignorance, that is, ignorance of what other people think (or are willing to say they think). It is implicit in these findings that people seem to want not to conform, but to be different from others in a desirable way. While highly suggestive, however, the “mere exposure” finding does not confirm the social influence account; it is possible that the views of others simply provide an informational signal, quite apart from arguments, and hence that people move not in order to maintain reputation, but to do what is right.

78. The phenomenon of group polarization may seem to raise doubts about rational actor models in economics or law. (On those models generally, see GARY S. BECKER, *ACCOUNTING FOR TASTES* (1996); *BEHAVIORAL LAW AND ECONOMICS* (Cass R. Sunstein ed., 2000).) But individual behavior within groups, as described thus far, should create no such doubts. It is certainly rational to make assessments on the basis of arguments offered; if the most numerous and convincing arguments seem to justify a shift, individual shifts are entirely rational. More difficult questions might seem to be raised by “social influence” accounts of group polarization. But it is also rational for people to care about their reputations. If they change their assessment because of reputational considerations, then maintaining a certain reputation is simply part of what people care about (and there is nothing irrational about that). If people shift not for reputational reasons but because of a certain self-conception—if, for example, they think of themselves as people who are bold, or committed to a strong national defense, or left of center on issues of race—a change in position, after exposure to the views of others, also seems entirely rational.

original tilt.⁷⁹

There is a related possibility, not quite reducible to either of the two standard arguments, but incorporating elements of each. In their individual judgments, people are averse to extremes; they tend to seek the middle of the relevant poles.⁸⁰ It is possible that when people are making judgments individually, they err on the side of caution, expressing a view in the direction that they really hold, but stating that view cautiously, for fear of seeming extreme. Once other people express supportive views, the relevant inhibition disappears, and people feel free to say what, in a sense, they really believe. There appears to be no direct test of this hypothesis, but it is reasonable to believe that the phenomenon plays a role in group polarization and choice shifts.⁸¹

D. *Refinements—and Depolarization*

I now turn to some refinements that complicate the basic account of group polarization. First, it matters a great deal whether people consider themselves part of the same social group as the other members. A sense of shared identity heightens the shift, and a belief that identity is not shared reduces and sometimes eliminates it. Second, deliberating groups tend to depolarize if they consist of equally opposed subgroups and if members have a degree of flexibility in their positions. Both of these findings have great relevance to any account of the relationship between group polarization and democratic institutions, as Part V shows.

1. *Statistical Regularities*

Of course, not all groups polarize. If the people defending the original tendency are particularly unpersuasive, group polarization is unlikely to

79. BROWN, *supra* note 8, at 219-20. An experimental study (on willingness to sacrifice money for the sake of fairness) has found support for the social influence explanation but not for the persuasive arguments explanation. Cason & Mui, *supra* note 8, at 1476-78.

80. See, e.g., Mark Kelman et al., *Context-Dependence in Legal Decision Making*, in BEHAVIORAL LAW AND ECONOMICS, *supra* note 78, at 61, 71-76 (discussing, inter alia, findings that the same option is evaluated more favorably when it is seen as the intermediate choice within a range of options).

81. A related possibility is that hearing other similar opinions produces greater confidence in individual positions, opening members to more extreme judgments in the same direction. This possibility has been raised by Heath and Gonzalez. Chip Heath & Rich Gonzalez, *Interaction with Others Increases Decision Confidence but Not Decision Quality: Evidence Against Information Collection Views of Interactive Decision Making*, 61 ORG. BEHAV. & HUM. DECISION PROCESSES 305, 318-19 (1995).

occur. If the outliers are especially convincing, groups may even shift away from their original tendency and in the direction held by few or even one.⁸²

Sometimes, moreover, external constraints or an external “shock” may prevent or blunt group polarization. Group members with well-defined views on a certain issue may be prone to polarization, but in order to maintain political effectiveness, even basic credibility, they will sometimes present a relatively moderate face, publicly or even privately. Groups that have started to polarize may move toward the middle in order to promote their own legitimacy or because of new revelations of one kind or another. Perhaps something like this happened with the Democratic Party in the early 1990s and the Republican Party in the late 1990s.

2. *Affective Factors and the Role of Confidence*

Affective factors are quite important in group decisions, and when manipulated, such factors will significantly increase or decrease polarization. If group members are linked by affective ties, dissent is significantly less frequent.⁸³ The existence of affective ties thus reduces the number of divergent arguments and also intensifies social influences on choice. Hence people are less likely to shift if the direction advocated is being pushed by unfriendly group members; the likelihood of a shift, and its likely size, are increased when people perceive fellow members as friendly, likeable, and similar to them.⁸⁴ In the same vein, physical spacing tends to reduce polarization; a sense of common fate and intragroup similarity tends to increase it, as does the introduction of a rival “outgroup.”⁸⁵ Mistakes are thus likely to be increased when group members are united mostly through

82. This is of course the plot of the movie *Twelve Angry Men*, in which the single holdout, played by Henry Fonda, shifts the judgment of the jury. See TWELVE ANGRY MEN (United Artists 1957); see also BROWN, *supra* note 8, at 229-39 (discussing the movie’s psychological realism).

83. Brooke Harrington, *The Pervasive Effects of Embeddedness in Organizations* 23-24 (1999) (unpublished manuscript, on file with *The Yale Law Journal*).

84. Hermann Brandstätter, *Social Emotions in Discussion Groups*, in DYNAMICS OF GROUP DECISIONS 93, 93-97, 106-08 (Hermann Brandstätter et al. eds., 1978). In TURNER ET AL., *supra* note 11, at 154-59, the authors attempt to use this evidence as a basis for a new synthesis, one that they call “a self-categorization theory of group polarization.” *Id.* at 154. In this account, “persuasion is dependent upon self-categorizations which create a common identity within a group” and polarization occurs “because group members adjust their opinion in line with their image of the group position (conform) and more extreme, already polarized, *prototypical* responses determine this image.” *Id.* at 156. When a group is leaning in a certain direction, the perceived “prototype” is determined by where the group is leaning, and this is where individuals will shift. *Id.* For possible differences in predictions, and supporting evidence, see *id.* at 158-70. An especially interesting implication, perhaps in some tension with the persuasive arguments theory, is that a group of comparative extremists will show a comparatively greater shift toward extremism. *Id.* at 158.

85. *Id.* at 151.

affective ties, and not through concentration on a particular task; it is in the former case that alternative views will less likely find expression.⁸⁶

The confidence of particular members also plays an important role.⁸⁷ Indeed, part of the reason for group polarization appears to be that as a class, extreme positions tend to be more confidently held. This point is an important complement to the persuasive arguments theory: The persuasiveness of arguments depends not simply on the grounds given, but also on the confidence with which they are articulated. Group polarization can also be fortified through “exit,” as members leave the group because they reject the direction in which things are heading. If exit is pervasive, the tendency to extremism will be aggravated.

3. *Identity and Solidarity*

If people think of themselves as part of a group having a degree of solidarity, group polarization is all the more likely, and it is likely to be more extreme.⁸⁸ Thus when the context emphasizes each person’s membership in the social group engaging in deliberation, polarization increases.⁸⁹ This finding is in line with more general evidence that social ties among deliberating group members tend to suppress dissent and in that way to lead to inferior decisions.⁹⁰ This should not be surprising. If ordinary findings of group polarization are a product of social influences and limited argument pools, it stands to reason that when group members think of one another as similar along a salient dimension, or if some external factor (for example, politics, geography, race, or sex) unites them, group polarization is heightened.⁹¹ Considerable evidence so suggests.⁹²

4. *Depolarization and Deliberation Without Shifts*

Is it possible to construct either groups that will depolarize—that will tend toward the middle—or groups whose members will not shift at all? Both phenomena seem to be real in actual deliberating bodies. In fact, the persuasive arguments theory implies, and evidence demonstrates, that there is depolarization if and when new persuasive arguments are offered that are

86. Harrington, *supra* note 35, at 26.

87. Maryla Zaleska, *The Stability of Extreme and Moderate Responses in Different Situations*, in *GROUP DECISION MAKING* 163, 164 (Hermann Brandstätter et al. eds., 1982).

88. WALLACE, *supra* note 3, at 73-76; Abrams et al., *supra* note 12, at 113-16; Spears et al., *supra* note 70, at 130-31.

89. Spears et al., *supra* note 70, at 122-24.

90. Harrington, *supra* note 83, at 26.

91. See the illuminating remarks on herd behavior, repression, and social identity in AMARTYA SEN, *REASON BEFORE IDENTITY* 19-22 (1999).

92. TURNER ET AL., *supra* note 11, at 159-62; Abrams et al., *supra* note 12, at 98-99; Spears et al., *supra* note 70, at 123-24, 130-31.

opposite to the direction initially favored by group members. Depolarization, rather than polarization, also occurs when a group consists of individuals drawn equally from two extremes.⁹³ If people who initially favor caution are put together with people who initially favor risk-taking, the group judgment will move toward the middle.

Group members with extreme but not fixed positions sometimes shift to a more moderate position.⁹⁴ Consider a study consisting of six-member groups specifically designed to contain two subgroups, each consisting of three people initially committed to opposed extremes.⁹⁵ The effect of discussion was to produce movement toward the center. One reason may be the existence of persuasive arguments in both directions.⁹⁶ A study of opposed subgroups found the greatest depolarization with obscure matters of fact (for example, how far below sea level is the town of Sodom)—and the least depolarization with highly visible public questions (for example, whether capital punishment is justified). Matters of personal taste depolarized a moderate amount (for example, preference for basketball or football, or for colors for painting a room).⁹⁷

These findings fit well with the persuasive arguments account of polarization. Note too that when people have a fixed view of some highly salient public issue, they are likely to have heard a wide range of arguments in various directions, producing a full argument pool, and additional discussion is not likely to produce movement. “[F]amiliar and long-debated issues do not depolarize easily.”⁹⁸ With respect to such issues, people are less likely to shift at all. And when one or more people in a group know the right answer to a factual question, the group is likely to shift in the direction of accuracy.⁹⁹

5. *Several Regularities*

These remarks suggest some general, commonsense conclusions about how and when group discussion will move pre-deliberation opinions. Views based on a great deal of thought are least likely to shift; depolarization can occur with equal subgroups tending in opposite directions; groups will usually shift in the direction of an accurate factual judgment where one or

93. Eugene Burnstein, *Persuasion as Argument Processing*, in *GROUP DECISION MAKING*, *supra* note 87, at 103, 107-11.

94. Duncan A. Ferguson & Neil Vidmar, *Effects of Group Discussion on Estimates of Culturally Appropriate Risk Levels*, 20 *J. PERSONALITY & SOC. PSYCHOL.* 436, 443-44 (1971).

95. BROWN, *supra* note 8, at 225.

96. Amiram Vinokur & Eugene Burnstein, *Effects of Partially Shared Persuasive Arguments on Group-Induced Shifts: A Group-Problem-Solving Approach*, 29 *J. PERSONALITY & SOC. PSYCHOL.* 305, 314 (1974).

97. Burnstein, *supra* note 93, at 111-12.

98. BROWN, *supra* note 8, at 226.

99. See Fishkin & Luskin, *supra* note 6, at 29-30.

more members knows the truth; where views are not firmly held, but where there is an initial predisposition, group polarization is the general rule. The effects of discussion are also likely to depend on members' perception of the group and of their relationship to it. If a group consists of "people," less polarization is likely than if it consists of "Republicans" or "defenders of the Second Amendment" or "opponents of American imperialism." Depolarization may well occur in groups with equal subgroups having opposite tendencies. But this is less likely and less pronounced if (1) subgroup members have fixed positions, and (2) subgroup members know that they are members of identifiable groups, and that their co-discussants are members of different identifiable groups.

An important general conclusion is that polarization will likely occur among heterogeneous as well as homogeneous groups, so long as there is a determinate initial predisposition among members, taken as a whole. The fact that some people oppose the dominant view will not prevent polarization from occurring. On the other hand, larger shifts are naturally to be expected within groups of like-minded people,¹⁰⁰ and as noted, depolarization is likely if the group is lacking a median predeliberation view in one direction or another, or if it consists of equally opposed subgroups whose members are willing to listen to one another.

A final question involves the effect of holding the median constant, but varying the "width" of the distribution of view within the deliberating group. In case *A*, for example, a group of six people, asked for their views on nuclear power on a scale of -5 ("strongly opposed") to +5 ("strongly in favor"), might have a predeliberation median of +2, with one at +1, one at +3, and four at +2. Here the prediction is that the group would come out at +3, and also that this would be the median of postdeliberation views, anonymously expressed. But compare case *B*, with a group with a predeliberation median also of +2 but with two at +4, two at +2, and two at 0. Because the median is the same, the predicted outcome is the same in case *B* as case *A*; but is this likely? I am aware of only one study that tests this question: the punitive damage study referred to above, which also contains the largest data set on the choice shift phenomenon, involving 3000 people and 500 deliberating groups.¹⁰¹ In this study, the shift

100. This is an oversimplification. What really matters is the predeliberation mean; the more extreme means will produce more extreme shifts, and a more extreme mean can exist within a heterogeneous group (say, a mean of +3, on a scale of -5 to +5, with a range of views) than in a group of like-minded people (say, a mean of +1, on the same scale, because every member is at +1).

101. Schkade et al., *supra* note 8. There is also a question about the effects, in experiments or in the real world over time, of small differences in initial views. Is it really the case, for example, that a group with five people opposed to affirmative action, and six in favor, will end up in a very different place from a group with six people opposed, and five in favor? Probably the best answer is that because the initial median is so close in the two cases, significant shifts over time are unlikely in experimental settings and probably in the real world.

depended on the median predeliberation judgment, and was not affected by wide or narrow spreads.¹⁰²

E. *Actual Deliberation Within Identifiable Groups: Iterated “Polarization Games”?*

Studies of group polarization involve one-shot experiments. What are the implications for people who meet with each other not once, but on a regular basis? The basic answer is that if participants engage in repeated discussions—if, for example, they meet each month, express views, and take votes—there should be repeated shifts toward, and past, the defined pole. If a group of citizens is thinking about genetic engineering of food, or the minimum wage, or the World Trade Organization, the consequence of their discussions, over time, should be to lead in quite extreme directions. In these iterated “polarization games,” deliberation over time should produce a situation in which individuals hold positions more extreme than those of any individual member before the series of deliberations began. In fact, the idea of iterated polarization games seems far more realistic than the processes studied in one-shot experiments.

There appears to be no study of such iterated polarization games. But the hypothesized outcomes are less fanciful than they might seem. In the jury study referred to above, deliberating groups frequently came up with punishment ratings, and with dollar awards, as high as or even higher than that of any individual, prior to deliberation.¹⁰³ And it is not difficult to think of real-world groups in which deliberation, over time, appears to shift both groups and individuals to positions that they could not possibly have accepted early on.¹⁰⁴ Iterated polarization games are an important real-world phenomenon.

But this raises two questions: Why and when do groups stop polarizing? Why and when do they terminate at a certain point, or even shift in the opposite direction? Nothing in the literature on group polarization adequately answers these questions.¹⁰⁵ But it is possible to speculate that polarization often ends or reverses as a result of some external shock—as, for example, when new members add new arguments, when the simple self-

102. *Id.* at 1156.

103. *Id.* at 1140.

104. For a relevant discussion of deliberating groups in the 1960s, see JAMES MILLER, “DEMOCRACY IS IN THE STREETS”: FROM PORT HURON TO THE SIEGE OF CHICAGO (1994), which traces increased radicalism and the spread of radical ideas; compare this with the discussion of “ethnification” in Eastern Europe in Timur Kuran, *Ethnic Norms and Their Transformation Through Reputational Cascades*, 27 J. LEGAL STUD. 623, 648-49 (1998). See *infra* notes 125-126 and accompanying text.

105. See TURNER ET AL., *supra* note 11, at 152 (suggesting that there is no clear answer to the question, “For what range of situations is polarization predicted?”).

interest of political leaders produces a shift in direction,¹⁰⁶ or when new changes in fact or value alter the perspectives and incentives of group members. Social cascades often change direction as a result of such external shocks, as through the release of new information;¹⁰⁷ the same processes seem to terminate or to reverse group polarization.

F. *Rhetorical Asymmetry and the “Severity Shift”: A Pervasive Phenomenon?*

The previously discussed study of punitive damages awards by juries found a striking pattern for dollar awards.¹⁰⁸ For *any* dollar award above zero, the general effect of deliberation was to increase awards above those of the median voter. This was a kind of “severity shift.” Dollar awards did not simply polarize; while higher awards increased dramatically, as compared to the median of predeliberation votes, lower awards increased as well. Both the original experiment and a follow-up experiment suggest that the severity shift is a product of a rhetorical asymmetry that favors, other things being equal, the person or persons urging higher awards.¹⁰⁹ The reason appears to be that with respect to dollar awards involving a corporate defendant, certain arguments—“we need to deter this kind of conduct,” “we need to send a powerful signal,” “we need to attract their attention”—tend to have comparatively greater weight.¹¹⁰

Undoubtedly there are many other contexts that give rise to rhetorical asymmetry, and undoubtedly the asymmetry can affect outcomes in democratic institutions, as it did in the jury study. Legislative judgments about criminal punishment may, for example, involve an asymmetry of exactly this kind; those favoring greater punishment for drug-related offenses appear to be at a systematic advantage over those favoring lesser punishment. In certain settings, those favoring lower taxes, or more aid for scholarship students, or greater funding for environmental protection, may have a similar rhetorical advantage. For present purposes, the point is that when there is an initial distribution of views in a certain direction, and when existing norms give a rhetorical advantage to a more extreme movement in that direction, quite extreme shifts can be expected.¹¹¹

106. See the discussion of legislators *infra* Subsection IV.E.4. Note, as inevitably noisy real-world examples, the centrist shift of the Democratic Party from 1980 to 1992 and the apparent centrist shift of the Republican Party from 1995 to 2000.

107. Hirschleifer, *supra* note 18, at 198-200 (discussing effects of public releases of information).

108. Schkade et al., *supra* note 8.

109. *Id.* at 1161-62.

110. *Id.*

111. Data from the study of punitive damages awards strongly support this speculation, with many deliberating juries producing dollar awards higher—and sometimes significantly higher—than the highest individual dollar award prior to deliberation. See *id.* at 1155-56. See also Cason &

IV. POLARIZATION AND DEMOCRACY

In this Part, I discuss evidence of group polarization in legal and political institutions, and I trace some implications of that evidence for participants in a deliberative democracy. My principal aim is descriptive, not normative. The relevant processes often create serious trouble for deliberative bodies, a point taken up in more detail in Parts V and VI.

A. *Polarizing Events and Polarization Entrepreneurs*

Survey evidence shows that dramatic events, such as the assassination of Martin Luther King and civil rights protests, tend to polarize attitudes, with both positive and negative attitudes increasing within demographic groups. Discussion often hardens attitudes toward outsiders and social change; thus “[p]roposals for establishment of a halfway house or a correctional facility have typically elicited private apprehensions which, after discussion, become polarized into overt paranoia and hostility.”¹¹² In fact, it is possible to imagine “professional polarizers,” or “polarization entrepreneurs,”¹¹³ that is, political activists who have as one of their goals the creation of spheres in which like-minded people can hear a particular point of view from one or more articulate people, and also participate, actually or vicariously, in a deliberative discussion in which that point of view becomes entrenched and strengthened. Those seeking to promote social reform might begin by promoting discussions among people who tend to favor the relevant reform.

As an example, consider the success of Lois Marie Gibbs, a Love Canal resident who became the principal force behind the national concern with abandoned hazardous waste dumps.¹¹⁴ Gibbs self-consciously engaged in efforts to mobilize citizens around that issue, partly by promoting discussions by like-minded people, first in small groups, then in larger

Mui, *supra* note 8, at 1476-78, which finds that, regardless of the initial distribution of views, groups systematically shift in the direction of greater willingness to sacrifice their material self-interest for the sake of fairness. Cason and Mui study choice shifts in the distinctive setting of the Dictator Game, in which a proposer is asked to divide a certain sum of money between herself and another subject, who must accept any allocation that the proposer selects. They find that for groups generally, the consequence of group discussion is to shift the proposed allocation in the direction of equality.

112. David G. Myers, *Polarizing Effects of Social Interaction*, in GROUP DECISION MAKING, *supra* note 87, at 125, 135.

113. Compare the discussion of availability entrepreneurs in Kuran & Sunstein, *supra* note 25, at 733, and of norm entrepreneurs in Cass R. Sunstein, *Social Norms and Social Roles*, 96 COLUM. L. REV. 903, 909 (1996); these terms emphasize different kinds of signaling by political leaders, whether self-interested or altruistic. The polarization entrepreneur, like the availability entrepreneur, shows an awareness of certain aspects of human psychology that are easily exploited to produce movement in preferred directions.

114. See LOIS MARIE GIBBS, LOVE CANAL: THE STORY CONTINUES . . . (1998).

ones.¹¹⁵ In fact, the environmental area is filled with leaders who took advantage of cascade-like processes and group polarization.¹¹⁶

Polarization is also likely to be produced by magazines with identifiable political convictions, such as the *American Prospect*, the *Weekly Standard*, the *New Republic*, and the *National Review*; by Pat Robertson and his special television programs devoted to his preferred causes; and by talk radio hosts with distinctive positions that are generally shared by their audiences. Because the results of group polarization cannot be evaluated in the abstract, nothing need be dishonorable in these efforts. What can be said, in the abstract, is that attempts to ensure discussion among people with similar predispositions may succeed in increasing the confidence of individual participants and also in moving them toward more extreme positions. Thus would-be social reformers do well to create forums, whether in person, over the air, in cyberspace, or in print, in which people with similar inclinations frequently speak with one another and can develop a clear sense of shared identity.

B. *Outgroups*

Group polarization has particular implications for insulated “outgroups” and, in the extreme case, for the treatment of conspiracies. Recall that polarization increases when group members identify themselves along some salient dimension and especially when the group is able to define itself by contrast to another group.¹¹⁷ Outgroups are, by definition, in this position of self-contrast to others. Excluded by choice or coercion from discussion with others, such groups may become polarized in quite extreme directions, often in part because of group polarization. In the midst of Communist rule, for example, the anticommunist underground was subject to polarization,¹¹⁸ sometimes undoubtedly for the better, but sometimes for worse. Acts of extremism (including murders and suicides) by outgroups are a possible result,¹¹⁹ especially if we consider the fact that extreme groups show comparatively greater polarization.¹²⁰ There is also likely to be some rhetorical asymmetry within such groups, so that arguments in a certain direction have the automatic upper hand.

The tendency toward polarization among outgroups helps explain special concern about hate speech, by which group antagonisms can be

115. *Id.* at 28-110.

116. See PENINA GLAZER & MYRON GLAZER, *THE ENVIRONMENTAL CRUSADERS* 6-15, 50-57, 171-73 (1998).

117. TURNER ET AL., *supra* note 11, at 154.

118. I owe the example to Wiktor Osiatynski.

119. Cf. ARONSON, *supra* note 17, at 242-43 (discussing the mass suicide at Jonestown).

120. TURNER ET AL., *supra* note 11, at 158, 167-70.

heightened, and it simultaneously raises some questions about the idea that certain group discussions produce “consciousness raising.” It is possible, at least, that the consequence of discussion is not only or mostly to raise consciousness (an ambiguous idea to be sure), but to produce group polarization in one direction or another—and at the same time to increase confidence in the position that has newly emerged.¹²¹ This does not mean that consciousness is never raised; undoubtedly group discussion can identify and clarify problems that were previously repressed, or understood as an individual rather than social product. But nothing of this sort is established by the mere fact that views have changed and coalesced, and are held, post-discussion, with a high degree of confidence.¹²²

An understanding of group polarization also casts light on the imposition of liability for criminal conspiracy, which can be added to the penalty for the substantive offense in most jurisdictions. It is tempting to think that this kind of “doubling up” is indefensible, a form of overkill. But if the act of conspiring leads people moderately disposed toward criminal behavior to be more than moderately disposed, precisely because they are conspiring together, it makes sense, on grounds of deterrence, to impose additional penalties. Some courts have come close to recognizing this point in discussing the imposition of distinct sanctions on conspiracies.¹²³

C. *Feuds, Ethnic and International Strife, and War*

Group polarization is likely to be at work in feuds of all kinds. One of the characteristic features of feuds is that members of feuding groups tend to talk only to one another, fueling and amplifying their outrage, and solidifying their impression of the relevant events. Informational and reputational forces are very much at work here, producing cascade effects, and group polarization can lead members to increasingly extreme positions. These effects are present within ethnic groups and even nations, notwithstanding the usually high degree of national heterogeneity. In America, sharp divergences between whites and African Americans, on particular salient events or more generally, can be explained by reference to group polarization.¹²⁴

Timur Kuran has explored the broader international phenomenon of “ethnification.” Kuran’s basic claim is that in many nations, including

121. Heath & Gonzalez, *supra* note 81, at 323-24.

122. *Id.* at 322.

123. *E.g.*, Callanan v. United States, 364 U.S. 587, 593-94 (1961) (“Concerted action . . . decreases the probability that the individuals involved will depart from their path of criminality.”). I am grateful to Dan Kahan for pressing this point. Recall that while actual burglars show a cautious shift in general conversation, their practices show a shift toward greater risk-taking (and correspondingly greater dangers for ordinary citizens). Cromwell et al., *supra* note 67.

124. See the treatment in JACOBS, *supra* note 14, at 140-51.

Turkey and the former Yugoslavia, ethnic strife is not the product of a reawakening of long-suppressed resentments, but instead of reputational cascades. In this process, a failure to engage in ethnically identified activity produces reputational sanctions, which grow in intensity over time as increasing numbers of people join the cascade. Initially people may be asked to dress in an ethnically identifiable way; later they may be asked to engage in certain celebrations and to participate in meetings; still later they may be asked to segregate themselves. Hence “*the fears and antagonisms that accompany high levels of ethnic activity may be a result of ethnification rather than its root cause.*”¹²⁵ Kuran does not refer to group polarization. But an understanding of this phenomenon would much fortify his analysis, by showing how within-group discussion (which is, under conditions of ethnification, an increasingly large percentage of total discussion) can ensure that ethnic groups, and individual members of ethnic groups, end up with a far stronger ethnic identification than the median member had before discussions began. In the extreme case, the result might be war.¹²⁶ And when a war begins, group polarization, if it operates at the national level, can help ensure continued hostility and antagonism.

D. *The Internet, Communications Policy, and Mass Deliberation*

Many people have expressed concern about processes of social influence on the mass media and the Internet.¹²⁷ The general problem is said to be one of fragmentation, with certain people hearing more and louder versions of their own preexisting commitments, thus reducing the benefits that come from exposure to competing views and unnoticed problems.¹²⁸ With greater specialization, people are increasingly able to avoid general interest newspapers and magazines and to make choices that reflect their own predispositions. The Internet makes it possible for people to design their own highly individual communications packages, filtering out troublesome issues and disfavored voices. Long before the Internet, it was possible to discuss the “racial stratification of the public sphere” by reference to divergences between white and African-American newspapers.¹²⁹ New communications technologies may increase this phenomenon.¹³⁰

125. Kuran, *supra* note 104, at 648.

126. *See id.* at 650-51.

127. *E.g.*, ANDREW L. SHAPIRO, THE CONTROL REVOLUTION 124-32 (1999).

128. *Id.* at 124-28.

129. JACOBS, *supra* note 14, at 144-45.

130. *Id.*

Group polarization explains why a fragmented communications market may create problems.¹³¹ A “plausible hypothesis is that the Internet-like setting is most likely to create a strong tendency toward group polarization when the members of the group feel some sense of group identity.”¹³² If certain people are deliberating with many like-minded others, views will not be reinforced, but instead will be shifted to more extreme points. This cannot be said to be bad by itself—perhaps the increased extremism is good—but it is certainly troublesome if diverse social groups are led, through predictable mechanisms, toward increasingly opposing and ever more extreme views. Processes of this general sort have threatened both peace and stability in some nations.¹³³ Both fragmentation and violence are possible results. Group polarization can intensify if people are speaking anonymously and if attention is drawn, through one or another means, to group membership. Many Internet discussion groups have precisely this feature.¹³⁴ The Internet may thus serve, for many, as a breeding ground for extremism.

An understanding of group polarization raises more general issues about communications policy. Under the “fairness doctrine,” now largely abandoned,¹³⁵ broadcasters were required to devote time to public issues and to allow an opportunity for opposing views to speak. The second prong of the doctrine was designed to ensure that listeners would not be exposed to any single view. When the FCC abandoned the fairness doctrine, it did so on the ground that this second prong often led broadcasters to avoid controversial issues entirely, and to present views in a way that suggested a bland uniformity.¹³⁶ Subsequent research has suggested that the elimination of the fairness doctrine has indeed produced a flowering of controversial substantive programming, frequently with an extreme view of one kind or another; consider talk radio.¹³⁷ Typically this is regarded as a story of wonderfully successful deregulation. But from the standpoint of group polarization, things are more complicated. The growth of issues-oriented programming with a strong, often extreme view may create group polarization, and all too many people might be exposed to louder echoes of their own voices, resulting in social fragmentation, enmity, and

131. A detailed discussion can be found in CASS R. SUNSTEIN, *REPUBLIC.COM* (draft of June 25, 2000) (forthcoming 2001) (manuscript on file with *The Yale Law Journal*).

132. WALLACE, *supra* note 3, at 78.

133. Kuran, *supra* note 104, at 635-51.

134. Compare the demonstration of serious errors within online working groups in Ross Hightower & Luftus Sayeed, *The Impact of Computer-Mediated Communication Systems on Biased Group Discussion*, 11 *COMPUTERS IN HUM. BEHAV.* 33 (1995).

135. *Syracuse Peace Council v. FCC*, 867 F.2d 654 (D.C. Cir. 1989).

136. *Id.* at 661.

137. Thomas W. Hazlett & David W. Sosa, *Was the Fairness Doctrine a “Chilling Effect”?* *Evidence from the Postderegulation Radio Market*, 26 *J. LEGAL STUD.* 279 (1997) (offering an affirmative answer to the question in the title).

misunderstanding. Perhaps it is better for people to hear fewer controversial views than for them to hear a single such view stated over and over again.

It is not clear what can be done about this situation. But it certainly makes sense to consider communications initiatives that would ensure that people are exposed to a range of reasonable views, not simply one. This was the original inspiration for the fairness doctrine, and there is reason to encourage media outlets to implement the same goal today. Thus Habermas's suggestion:

The diffusion of information and points of view . . . is not the only thing that matters in public processes of communication, nor is it the most important. . . . [T]he rules of a *shared* practice of communication are of greater significance for structuring public opinion. Agreement on issues and contributions *develops* only as the result of more or less exhaustive controversy in which proposals, information, and reasons can be more or less rationally dealt with.¹³⁸

Perhaps a code of fair programming could promote voluntary self-regulation in this direction.¹³⁹ With respect to the Internet, Andrew Shapiro has suggested public subsidy of a civic icon that would promote exposure to substantive discussions from a variety of viewpoints.¹⁴⁰ An appreciation of group polarization suggests the need for creative approaches designed to ensure that people do not simply read their "Daily Me."¹⁴¹

E. *Deliberation and Polarization in Public Institutions*

I now turn to brief discussions of group polarization in government institutions.

1. *Juries*

Group polarization on juries is well-documented; the jury, in fact, is the only legal institution for which direct evidence of group polarization exists. In experimental settings, polarization has been found in numerous instances with respect to guilt and innocence, and indeed this appears to be an uncontradicted finding.¹⁴² Outside of the experimental setting, we know that

138. HABERMAS, *supra* note 5, at 362.

139. Cass R. Sunstein, *Television and the Public Interest*, 88 CAL. L. REV. 499, 549-57 (2000).

140. SHAPIRO, *supra* note 127, at 205-07; *see also* SUNSTEIN, *supra* note 131 (manuscript at 102-14) (discussing Shapiro's proposal and other reforms designed to promote exposure to diverse ideas).

141. The term comes from Nicholas Negroponte. *See* SHAPIRO, *supra* note 127, at 45.

142. BROWN, *supra* note 8, at 227-29 (collecting studies).

the median predeliberation verdict predicts the final outcome ninety percent of the time in cases where juries do not hang; this provides “powerful presumptive evidence that group polarization occurs in real juries.”¹⁴³

2. *Independent Regulatory Commissions*

The twentieth century has seen the rise of a number of “independent” regulatory commissions, including the Federal Communications Commission, the Federal Trade Commission, and the National Labor Relations Board. A striking but generally overlooked provision of the relevant statutes requires bipartisan membership: The independent commissions cannot have more than a certain number of members from any one political party.¹⁴⁴ A simple and undoubtedly correct explanation of this unusual requirement is that Congress wanted to ensure that no commission would be dominated by any single party. But an understanding of group polarization would strengthen any such concern on Congress’s part. An independent agency that is all Democratic, or all Republican, might polarize toward an extreme position, likely more extreme than that of the median Democrat or Republican, and possibly more extreme than that of any member standing alone. A requirement of bipartisan membership can operate as a check against movements of this kind.

3. *Multimember Courts*

Group polarization might also occur on multimember courts. Notwithstanding platitudes about judicial neutrality, judges often have a great deal of latitude, sometimes in the ultimate outcome, more often in determining the reach of their decision. If a court consists of three or more like-minded judges, it may well end up with a relatively extreme position, more extreme in fact than the position it would occupy if it consisted of two like-minded individuals and one of a different orientation.

There is little direct confirmation of this general proposition. But considerable support comes from two studies of judicial behavior on the U.S. Court of Appeals for the District of Columbia Circuit.¹⁴⁵ The first study finds a tendency toward more extreme results when a panel consists of judges from a single political party; most notably, a panel of three Republican judges is far more likely than a panel of two Republicans and

143. *Id.* at 229. See Schkade et al., *supra* note 8, for a treatment of choice shifts on juries deliberating about punitive damages.

144. 47 U.S.C. § 154(b)(5) (1994) (FCC); 15 U.S.C. § 41 (1994) (FTC).

145. Frank B. Cross & Emerson H. Tiller, *Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals*, 107 YALE L.J. 2155 (1998); Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717, 1755 (1997).

one Democrat to reverse an environmental decision at the behest of industry challengers.¹⁴⁶ This is precisely what would be predicted by group polarization: A homogeneous panel is far more likely to go in an extreme direction than a mixed panel.

The second study is a bit more complex.¹⁴⁷ Under *Chevron v. NRDC*,¹⁴⁸ courts are supposed to uphold agency interpretations of law so long as the interpretations are “reasonable.” The study found a dramatic difference in the way politically divided panels (with judges appointed by presidents from more than one party) and “unified” panels (with judges appointed by presidents from only one party) on the D.C. Circuit interpreted this stricture. On *divided* panels in which a majority of the panel might be expected to be hostile to the agency on, broadly speaking, political grounds, the court deferred to the agency 62% of the time.¹⁴⁹ But on *unified* panels, in which the court might be expected to be hostile to the agency, the court upheld the agency interpretation only 33% of the time.¹⁵⁰ Note that this was the only asymmetry in the data; when courts were expected to uphold the agency’s decision on political controls, they did so over 70% of the time, whether unified (71% of the time) or divided (84% of the time).¹⁵¹ There is no smoking gun here, but it seems reasonable to speculate that the seemingly bizarre result—a mere 33% validation rate in cases in which the panel was unified—reflects a process of group polarization. A group of like-minded judges may well take the relatively unusual step of rejecting an agency interpretation, whereas a divided panel, with a check on any tendency toward extreme outcomes, is more likely to take the conventional route.

4. Legislatures

Legislators are likely to be susceptible to group polarization, partly because of the effects of limited argument pools, and perhaps above all because of social influence (including the importance of conveying a proper signal to fellow legislators and constituents). Imagine, for example, that a group of Republicans and a group of Democrats are thinking about how to vote on a proposed law. If Republicans are speaking mostly with Republicans, and if Democrats are speaking mostly with Democrats, one should expect a hardening of views toward the more extreme points. Undoubtedly this is part (certainly not all) of the explanation of party-line

146. Revesz, *supra* note 145, at 1755.

147. See Cross & Tiller, *supra* note 145.

148. 467 U.S. 837 (1984).

149. Cross & Tiller, *supra* note 145, at 2172.

150. *Id.*

151. *Id.*

voting. Mechanisms of group polarization will sometimes be at work with constituents as well. It is easy to imagine a society in which Republicans speak mostly with each other; one can imagine a society in which Democrats speak mostly with one another too. If this is the situation, polarization should occur within political camps.

V. DELIBERATIVE TROUBLE?

This Part turns to normative issues involving the relationship among group polarization, democratic theory, and legal institutions. I focus in particular on the implications of group polarization for institutional design, with special reference to the uses of heterogeneity and the complex issues presented by deliberation inside particular “enclaves.”

The central problem is that widespread error and social fragmentation are likely to result when like-minded people, insulated from others, move in extreme directions simply because of limited argument pools and parochial influences. As an extreme example, consider a system of one-party domination, which stifles dissent in part because it refuses to establish space for the emergence of divergent positions; in this way, it intensifies polarization within the party while also disabling external criticism.

In terms of institutional design, the most natural response is to ensure that members of deliberating groups, whether small or large, will not isolate themselves from competing views—a point with implications for multimember courts, open primaries, freedom of association, and the architecture of the Internet. Here, then, is a plea for ensuring that deliberation occurs within a large and heterogeneous public sphere, and for guarding against a situation in which like-minded people wall themselves off from alternative perspectives.

But there is a difficulty with this response: A certain measure of isolation will, in some cases, be crucial to the development of ideas and approaches that would not otherwise emerge and that deserve a social hearing.¹⁵² Members of low-status groups are often quiet within heterogeneous bodies, and deliberation in such bodies tends to be dominated by high-status members. Any shift—in technology, norms, or legal practice—that increases the number of deliberating enclaves will likewise increase the diversity of society’s aggregate “argument pool” while also increasing the danger of extremism and instability, ultimately even violence. Shifts toward a general “public sphere,” without much in the way of enclave deliberation, will decrease the likelihood of extremism and instability, but at the same time produce what may be a stifling

152. See the empirical references in Christensen & Abbott, *supra* note 16, at 273-77.

uniformity.¹⁵³ And shifts toward more in the way of enclave deliberation will increase society's aggregate "argument pool," and hence enrich the marketplace of ideas,¹⁵⁴ while also increasing fragmentation, hostility, and perhaps even violence.

No algorithm is available to solve the resulting conundrums. But some general lessons do emerge. It is important to ensure social spaces for deliberation by like-minded persons, but it is equally important to ensure that members of the relevant groups are not isolated from conversation with people having quite different views. The goal of that conversation is to promote the interests of those inside and outside the relevant enclaves, by subjecting group members to competing positions, by allowing them to exchange views with others and to see things from their point of view, and by ensuring that the wider society does not marginalize, and thus insulate itself from, views that may turn out to be right or at least informative.

Ideas of this kind have a central place in the constitutional framework, with the system of checks and balances, bicameralism, and the Framers' explicit rejection of the "right to instruct" representatives.¹⁵⁵ They also have an important emerging place on the Internet, where there are many current proposals to create a "public sphere" involving discussions among diverse people who would otherwise be in contact largely with those who are like-minded.¹⁵⁶ In these ways, an understanding of group polarization helps provide a better sense of how institutions, both old and new, might be restructured in the service of democratic ideals.

A. *Doubts and Questions*

1. *Why Deliberate?*

It should be clear that the phenomenon of group polarization, placed alongside the phenomenon of social cascades, raises severe doubts about the view that deliberation is a simple or unambiguous good. Of course deliberation might be justified on many different grounds.¹⁵⁷ It may be that on the question at issue, there is a truth of the matter—a correct answer—and deliberation might be thought to be the best way of reaching it. Perhaps group decisions, based on an exchange of reasons, are more likely to be right than decisions made by individuals. Alternatively, society might favor

153. See the discussion of the shift, over time, from the "unthinkable" to the "unthought" in KURAN, *supra* note 30, at 176-95.

154. See the discussion of African-American newspapers in JACOBS, *supra* note 14, at 31-53.

155. *Infra* text accompanying notes 163-172.

156. *E.g.*, SHAPIRO, *supra* note 127, at 205-07; SUNSTEIN, *supra* note 131 (manuscript at 102-03, 111-14).

157. For various perspectives, see DELIBERATIVE DEMOCRACY, *supra* note 5.

deliberation on a quite different ground: Doubting whether there is a truth of the matter, a society might seek a deliberative process on the theory that this is the only reasonable and fair way to reach a decision that will be imposed on the group.¹⁵⁸ Or deliberation might be seen as a way to reach some sort of shared understanding or to ensure a form of mutual respect. Group polarization does not create obvious difficulties for all of these accounts. But it does raise real questions about the widespread idea that deliberation is the best way of producing right answers.

If the effect of deliberation is to move people toward a more extreme point in the direction of their original tendency, why is it anything to celebrate? The underlying mechanisms do not provide much reason for confidence. If people shift their position in order to maintain their reputation and self-conception, is there any reason to think that deliberation is making things better rather than worse? If shifts occur as a result of partial and frequently skewed argument pools, the results of deliberative judgments may be far worse than the results of simply taking the median of predeliberation judgments.

To be sure, those who emphasize the ideals associated with deliberative democracy tend to emphasize its preconditions, which include political equality, an absence of strategic behavior, full information, and the goal of “reaching understanding.”¹⁵⁹ In real-world deliberations, behavior is often strategic, and equality is often absent in one or another form. But group polarization is likely to occur even in the face of equality and of entirely conscientious efforts at reaching both truth and understanding. The existence of a limited argument pool, strengthening the existing tendency within the group, will operate in favor of group polarization even if no individual behaves strategically. By itself, this will produce group polarization, whether or not social influence is operating.

In any case, social influences need not be inconsistent with the effort to produce truth and understanding. When people attempt to position themselves in a way that fits with their best self-conception or their preferred self-presentation, nothing has gone wrong, even from the standpoint of deliberation’s most enthusiastic defenders.¹⁶⁰ Perhaps group polarization could be reduced or even eliminated if we emphasized that full

158. See the discussion of imperfect procedural justice and pure procedural justice in RAWLS, *supra* note 2, at 83-90.

159. *E.g.*, 1 JÜRGEN HABERMAS, *THE THEORY OF COMMUNICATIVE ACTION* 99 (Thomas McCarthy trans., Beacon Press 1984) (1981). Thus Habermas distinguishes between strategic and communicative action and stresses “the cooperatively pursued goal of reaching understanding.” *Id.* Compare Habermas’s view with the treatment in GUTMANN & THOMPSON, *supra* note 5, at 52-94, referring to the idea of reciprocity, which emphasizes the desire to justify one’s position by reference to reasons.

160. See Robert E. Goodin, *Laundering Preferences*, in *FOUNDATIONS OF SOCIAL CHOICE THEORY* 75, 77-90 (Jon Elster & Aanund Hylland eds., 1986) (discussing self-censorship as a check on invidious views).

information is a precondition for good deliberation. By hypothesis, argument pools would not be limited if all information were available. But that requirement is extremely stringent, and if there is already full information, the role of deliberation is greatly reduced.¹⁶¹ In any case, the group polarization phenomenon suggests that, in real-world situations, deliberation is hardly guaranteed to increase the likelihood of arriving at truth.

2. *Movements Right and Wrong*

Polarization does not necessarily mean that there has been a movement in the wrong direction. Perhaps the more extreme tendency is better. Recall that group polarization is likely to have fueled the antislavery movement and many other movements that deserve widespread approval. In the context of punitive damages awards, perhaps a severity shift produces good outcomes. Extremism need hardly be a word of opprobrium; everything depends on what extremists are arguing *for*. But when group discussion leads people to more strongly held versions of the same view with which they began, and when social influences and limited argument pools are responsible, there is little reason for great confidence in the effects of deliberation.

If it is possible to identify a particular viewpoint as unreasonable, it is also possible to worry about group discussion among people who share that viewpoint. As noted, this is a basis for justifying the decision to make criminal conspiracy an independent crime; the act of conspiring itself raises the stakes, and by hypothesis the actions ultimately at issue are unlawful and beyond the pale.¹⁶² An analogy can be found in discussions within hate groups, on the Internet and elsewhere. If the underlying views are unreasonable, it makes sense to fear that these discussions may fuel increasing hatred. This does not mean that the discussions can or should be regulated in a system dedicated to freedom of speech. But it does raise questions about the idea that “more speech” is necessarily an adequate remedy.

B. *The Virtues of Heterogeneity*

The simplest lesson here involves both individual susceptibility and institutional design. For many people, mere awareness of the role of limited argument pools and social influences might provide some inoculation

161. The role of deliberation is not eliminated. There remains the question of what to do, given a certain understanding of the facts.

162. *Supra* note 123 and accompanying text.

against inadequately justified movements of opinion within groups. More importantly, institutions should be designed to ensure that when shifts are occurring, it is not because of arbitrary or illegitimate constraints on the available range of arguments. This is a central task of constitutional design. In this light, a system of checks and balances might be explained, not as an undemocratic check on the will of the people, but as an effort to protect against potentially harmful consequences of group discussion.¹⁶³

We have seen that the system of bicameralism, often challenged on populist grounds,¹⁶⁴ might be defended by reference to the risk of group polarization. Indeed, James Wilson's great lectures on law spoke of bicameralism in these terms, referring to "instances, in which the people have become the miserable victims of passions, operating on their government without restraint," and seeing a "single legislature" as prone to "sudden and violent fits of despotism, injustice, and cruelty."¹⁶⁵ Efforts to assure a plurality of views on regulatory commissions and courts can be defended on similar grounds. As supporting evidence, consider the findings that cohesive groups of like-minded people whose members are connected by close social ties often suppress dissent and reach inferior decisions, whereas heterogeneous groups, building identification through focus on a common task rather than through other social ties, tend to produce the best outcomes.¹⁶⁶

C. *A Thought Experiment: Vindicating Hamilton*

To explore some of the advantages of heterogeneity, imagine a deliberating body consisting of all citizens in the relevant group; this may mean all citizens in a community, a state, a nation, or the world. By hypothesis, the argument pool would be very large. It would be limited only to the extent that the set of citizen views was similarly limited. Social influences would undoubtedly remain. Hence people might shift because they want to maintain their reputation and self-conception, and to do this they must stand in a certain relation to the rest of the group. But to the extent that deliberation revealed to people that their private position was different from what they thought it was in relation to the group, any shift would be in response to an accurate understanding of all relevant citizens, and not a product of a skewed group sample.

163. *Federalist No. 51* can well be read in this light. THE FEDERALIST NO. 51 (James Madison).

164. See the recent proposals in Minnesota, outlined and discussed in George Will, Editorial, *One-House Town*, WASH. POST, Nov. 4, 1999, at A35, 1999 WL 23312778.

165. JAMES WILSON, *Lectures on Law*, in 1 THE WORKS OF JAMES WILSON 291 (Robert Green McCloskey ed., 1967).

166. Harrington, *supra* note 83, at 25-26.

This thought experiment does not suggest that the hypothesized deliberating body would be ideal. Perhaps all citizens, presenting all individual views, would offer a skewed picture from the normative point of view; in a pervasively unjust society, a deliberating body consisting of everyone may produce nothing to celebrate. Perhaps weak arguments would be made and repeated and repeated again, while good arguments would be offered infrequently. As I suggest below,¹⁶⁷ it is often important to ensure that there are enclaves in which polarization will take place, precisely in order to ensure the emergence of views that are suppressed, by social influences or otherwise, but are reasonable or even right. But at least a deliberating body of all citizens would remove some of the distortions in the group polarization experiments, where generally like-minded people, not exposed to others, shift in large part because of that limited exposure.¹⁶⁸ Many studies show that organizational performance is impaired by people's failure to voice diverse views,¹⁶⁹ and this is a tendency against which heterogeneity might guard.

Here, in fact, is a way of vindicating the passages from Hamilton and Rawls with which I began. On this view, Hamilton and Rawls are not naive enthusiasts for deliberation, oblivious to empirical realities, but are insisting on the advantages of heterogeneity and of a wide argument pool being placed before the deliberators. Indeed, this was Hamilton's claim with respect to the "jarring of parties" within a bicameral legislature, a process that, Hamilton contended, would "check the excesses of the majority"—excesses that can be reinterpreted in terms of the phenomena I have been describing. The Framers of the Bill of Rights originally rejected a "right to instruct," by which constituents could tell their representatives how to vote; the rejection was based on the idea that one job of the representative was to "consult" with people from different states of the Union, and to make decisions only after that consultation.¹⁷⁰ The rejection of the right to instruct, on the theory that deliberators should be talking to people with different experiences and viewpoints, seems to show an appreciation of the risks of group polarization.

As I have suggested, central features of the constitutional design, including the system of checks and balances, can be understood in similar terms. Indeed, Madison defended his preference for large election districts and long terms of service as a way of counteracting polarization-type forces

167. *Infra* text accompanying notes 173-177.

168. See Harrington, *supra* note 83, for supporting evidence in the context of concrete group tasks (involving investment clubs).

169. Charlan Jeanne Nemeth, *Differential Contributions of Majority and Minority Influence*, 93 PSYCHOL. REV. 23 (1986); Harrington, *supra* note 83.

170. SUNSTEIN, *supra* note 5, at 22.

within particular constituencies.¹⁷¹ Rawls's reference to the need to "combin[e] information and enlarg[e] the range of arguments" through discussion with a range of people strikes precisely the same note.¹⁷²

VI. ENCLAVE DELIBERATION AND SUPPRESSED VOICES

There are also potential vices in heterogeneity and potentially desirable effects from deliberating "enclaves," consisting of groups of like-minded individuals. It seems obvious that such groups can be extremely important in a heterogeneous society, not least because members of some demographic groups tend to be especially quiet when participating in broader deliberative bodies. A special advantage of "enclave deliberation" is that it promotes the development of positions that would otherwise be invisible, silenced, or squelched in general debate.¹⁷³ In numerous contexts, this is a great advantage; many social movements have been made possible through this route (as possible examples, consider feminism,¹⁷⁴ the civil rights movement, religious conservatism, environmentalism, and the movement for gay and lesbian rights).

The efforts of marginalized groups to exclude outsiders, and even of political parties to limit their primaries to party members, can be justified in similar terms. Even if group polarization is at work—perhaps *because* group polarization is at work—enclaves can provide a wide range of social benefits, not least because they greatly enrich the social argument pool. In fact, the First Amendment right of expressive association should be understood in precisely these terms. According protection to collective effort on behalf of shared goals is especially important in preserving political and cultural diversity and in shielding dissident expression from suppression by the majority.

A. *Low Status and High Status*

In deliberating bodies, high-status members tend to initiate communication more than others, and their ideas are more influential, partly

171. Cass R. Sunstein, *Interest Groups in American Public Law*, 38 STAN. L. REV. 29, 41-42 (1985).

172. RAWLS, *supra* note 2, at 359.

173. See, for example, the discussion of consciousness raising in CATHARINE A. MACKINNON, *TOWARD A FEMINIST THEORY OF THE STATE* 83-105 (1989).

174. A controversial and highly publicized case in point is the effort by the theologian Mary Daly to exclude men from her class at Boston University. See Carey Goldberg, *Facing Forced Retirement, Iconoclastic Professor Keeps On Fighting*, N.Y. TIMES, Aug. 15, 1999, at A13, 1999 WL 30476861; see also *Boy Scouts of America v. Dale*, 120 S. Ct. 2446 (2000) (upholding the associational interest and right of exclusion of Boy Scouts); *Roberts v. U.S. Jaycees*, 468 U.S. 609, 622 (1984) (rejecting the right of exclusion of the Jaycees).

because low-status members lack confidence in their own abilities, partly because they fear retribution.¹⁷⁵ For example, women's ideas are often less influential and sometimes are "suppressed altogether in mixed-gender groups,"¹⁷⁶ and in ordinary circumstances, cultural minorities have disproportionately little influence on decisions by culturally mixed groups.¹⁷⁷ It makes sense to promote deliberating enclaves in which members of multiple groups may speak with one another and develop their views.

But there is a serious danger in such enclaves. The danger is that through the mechanisms discussed here, members will move to positions that lack merit but are predictable consequences of the particular circumstances of enclave deliberation. In the extreme case, enclave deliberation may even put social stability at risk (for better or for worse). And it is impossible to say, in the abstract, that those who sort themselves into enclaves will generally move in a direction that is desirable for society at large or even for its own members. It is easy to think of examples to the contrary, such as the rise of Nazism, hate groups, and numerous "cults" of various sorts;¹⁷⁸ readers can think of their own preferred illustrations.

There is no simple solution to the dangers of enclave deliberation. Sometimes the threat to social stability is desirable. As Jefferson wrote, turbulence can be "productive of good. It prevents the degeneracy of government, and nourishes a general attention to . . . public affairs. I hold . . . that a little rebellion now and then is a good thing . . ." ¹⁷⁹ Any judgments about enclave deliberation are hard to make without a sense of the underlying substance—of what it is that divides the enclave from the rest of society. From the standpoint of institutional design, the problem is that any effort to promote enclave deliberation will ensure group polarization within a wide range of groups, some necessary to the pursuit of justice, others likely to promote injustice, and some potentially quite dangerous. This makes clearer the sense in which Edmund Burke's conception of representation—rejecting "local purposes" and "local

175. See Christensen & Abbott, *supra* note 16, at 273.

176. *Id.* at 274.

177. Catherine Kirchmeyer & Aaron Cohen, *Multicultural Groups: Their Performance and Reactions with Constructive Conflict*, 17 *GROUP & ORG. MGMT.* 153, 166 (1992). Interestingly, there is evidence that with changes in gender norms, some tasks show no gender differences in influence on groups. See Katherine W. Hawkins, *Effects of Gender and Communication Content on Leadership Emergence in Small Task-Oriented Groups*, 26 *SMALL GROUP RES.* 234, 243-44 (1995).

178. See, e.g., ARONSON, *supra* note 17, at 242-43 (discussing the role of cult leaders); LEON FESTINGER ET AL., *WHEN PROPHECY FAILS* 3-30 (1956) (discussing conditions for continued commitment to implausible beliefs, including the need for social support).

179. Letter from Thomas Jefferson to James Madison (Jan. 30, 1787), in *THE PORTABLE THOMAS JEFFERSON* 416-17 (Merrill D. Peterson ed., 1975).

prejudices” in favor of “the general reason of the whole”¹⁸⁰—is not contingently but instead *essentially* conservative (speaking purely descriptively, as a safeguard of existing practices). The reason is that the submersion of “local purposes” and “local prejudices” into a heterogeneous “deliberative assembly”¹⁸¹ will inevitably tend to weaken the resolve of groups—and particularly low-status or marginalized groups—whose purely internal deliberations would produce a high degree of polarization.

Hence James Madison—with his fear of popular passions producing “a rage for paper money, for an abolition of debts, for an equal division of property, or for any other improper or wicked project”¹⁸²—would naturally be drawn to a Burkean conception of representation, favoring large election districts and long length of service¹⁸³ to counteract the forces of polarization. By contrast, those who believe that “destabilization” is an intrinsic good,¹⁸⁴ or that the status quo contains sufficient injustice that it is worthwhile to incur the risks of encouraging polarization on the part of diverse groups, will, or should, be drawn to a system that enthusiastically promotes insular deliberation within enclaves.

In a nation in which most people are confused or evil, enclave deliberation may be the only way to develop a sense of clarity or justice, at least for some. But even in such a nation, enclave deliberation is unlikely to produce change unless the members of different enclaves are eventually brought into contact with others. In democratic societies, the best response is to ensure that any such enclaves are not walled off from competing views, and that at certain points, there is an exchange of views between enclave members and those who disagree with them. It is total or near-total self-insulation, rather than group deliberation as such, that carries with it the most serious dangers, often in the highly unfortunate (and sometimes deadly) combination of extremism with marginality.

B. *The Public Sphere and Appropriate Heterogeneity*

1. *The Public Sphere*

For a designer or leader of any institution, it makes sense to promote ample social space both for enclave deliberation and for discussions

180. Edmund Burke, Speech to the Electors (Nov. 3, 1774), in BURKE’S POLITICS: SELECTED WRITINGS AND SPEECHES OF EDMUND BURKE ON REFORM, REVOLUTION, AND WAR 116 (Ross J.S. Hoffman & Paul Levack eds., 1949).

181. *Id.*

182. THE FEDERALIST NO. 10, at 84 (James Madison) (Clinton Rossiter ed., 1961).

183. See Sunstein, *supra* note 171, at 41-42.

184. This is a possible reading of ROBERTO MANGABEIRA UNGER, THE CRITICAL LEGAL STUDIES MOVEMENT 39, 43, 53-57 (1986), which asserts the need for destabilization rights.

involving a broad array of views, including views of those who have been within diverse enclaves.¹⁸⁵ The idea of a “public sphere,” developed most prominently by Jürgen Habermas, can be understood as an effort to ensure a domain in which multiple views can be heard by people with multiple perspectives.¹⁸⁶ This understanding strongly supports current initiatives designed to ensure deliberation among dissimilar people on the Internet.¹⁸⁷ The fact of group polarization suggests that it could be desirable to take steps to reduce the likelihood that panels on federal courts of appeals consist solely of appointees of presidents of any single political party.¹⁸⁸

Of course, any argument pool will be limited. No one has time to listen to every point of view. But an understanding of group polarization helps show that heterogeneous groups are often a far better source of good judgments, simply because more arguments will be made available.

2. *A New Look at Group Representation*

An understanding of group polarization is also relevant to the continuing debate over group representation.¹⁸⁹ The central issue here is whether identifiable groups should be represented as such in political institutions. Perhaps political groups should be allowed to have representation to the extent that they are able to get more than a minimal share of the vote. Perhaps steps should be taken to increase the likelihood that members of disadvantaged or marginal groups have their own representatives in the deliberating body. An understanding of group polarization is hardly sufficient to reach a definitive conclusion about these issues. But at least it can be said that group representation should help counteract the risks of polarization, and susceptibility to cascade effects, that come from deliberation among like-minded people. At the same time, group representation should help reduce the dangers that come from insulation of those in the smaller enclave, by subjecting enclave representatives to a broader debate.

Seen in this light, the point of group representation is to promote a process in which those in the enclave hear what others have to say, and in which those in other enclaves, or in no enclaves at all, are able to listen to

185. See IRVING L. JANIS, *GROUPTHINK: PSYCHOLOGICAL STUDIES OF POLICY DECISIONS AND FIASCOS* 267-71 (2d ed. 1982).

186. See HABERMAS, *supra* note 9, at 231-50.

187. See SUNSTEIN, *supra* note 131 (manuscript at 102-03, 111-14).

188. See *supra* Subsection IV.E.3.

189. See, e.g., IRIS MARION YOUNG, *JUSTICE AND THE POLITICS OF DIFFERENCE* 183-91 (1990) (urging group representation as a means of ensuring expression of plural perspectives); Cass R. Sunstein, *Beyond the Republican Revival*, 97 *YALE L.J.* 1539, 1585-89 (1988) (urging group representation as a means of ensuring good deliberation about the public good). For an overview of the debate, see also the essays in *CHOOSING AN ELECTORAL SYSTEM: ISSUES AND ALTERNATIVES* (Arend Lijphart & Bernard Grofman eds., 1984).

people with very different points of view. When groups are represented as such in private and public institutions, it is to promote the airing of points of view that otherwise go unheard. And there is a further benefit to group representation: It works to prevent self-insulation on the part of group members by ensuring that they hear what others have to say as well. Here the benefits of group representation extend to both members and nonmembers of the relevant groups.

3. *Appropriate Heterogeneity*

For a deliberative democracy, a central question is how to ensure *appropriate* heterogeneity. For example, it would not make sense to say that in a deliberating group attempting to think through issues of affirmative action, it is important to allow exposure to people who think that slavery was good and should be restored. The constraints of time and attention call for limits to heterogeneity; and—a separate point—for good deliberation to take place, some views are properly placed off the table, simply because time is limited and they are so invidious, or implausible, or both. This point might seem to create a final conundrum: To know what points of view should be represented in any group deliberation, it is important to have a good sense of the substantive issues involved, indeed a sense sufficient to generate judgments about what points of view must be included and excluded. But if we already know that, does deliberation have any point at all?

The answer is that we often do know enough to see which views count as reasonable, without knowing which view counts as right, and this point is sufficient to allow people to construct deliberative processes that should correct for the most serious problems potentially created by group polarization. What is necessary is not to allow every view to be heard, but to ensure that no single view is so widely heard, and reinforced, that people are unable to engage in critical evaluation of the reasonable competitors.

Of course, the provision of diverse views does not guarantee good deliberation. Among other things, most people are subject to “confirmation bias,” in accordance with which exposure to a competing position will not dislodge and may even strengthen the antecedently held position.¹⁹⁰ On questions of morality and fairness, and undoubtedly other questions as well, those who listen to diverse opinions may well emerge from the experience with an enhanced belief in the soundness of their original commitment.¹⁹¹ But this is not a universal phenomenon, and at least an understanding of

190. Charles G. Lord et al., *Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence*, 37 J. PERSONALITY & SOC. PSYCHOL. 2098, 2102-04 (1979).

191. *Id.*

competing views is likely to weaken the forms of fragmentation and misunderstanding that come from deliberation among the like-minded.¹⁹²

C. *The Deliberative Opinion Poll: A Contrast*

James Fishkin has pioneered the idea of “deliberative polling,” in which small groups consisting of highly diverse individuals are asked to come together and to deliberate about various issues.¹⁹³ Deliberative opinion polls have now been conducted in several nations, including the United States, England, and Australia. Fishkin finds some noteworthy shifts in individual views, but he does not find a systematic tendency toward polarization. In his studies, individuals shift both toward and away from the median of predeliberation views. For present purposes, what is noteworthy about Fishkin’s experiments is that they do not involve isolated enclaves; this is part of the reason for the absence of polarization effects.

In England, for example, deliberation led to reduced interest in using imprisonment as a tool for combating crime.¹⁹⁴ The percentage believing that “sending 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 reduced from 78% to 65%.¹⁹⁵ Similar shifts were shown in the direction of greater enthusiasm for procedural rights of defendants and increased willingness to explore alternatives to prison.¹⁹⁶ 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%).¹⁹⁷ On many issues, deliberation increased the intensity with which people held their preexisting convictions.¹⁹⁸ These findings are consistent with the prediction of group polarization. But this was hardly a uniform pattern, and on some questions, deliberation increased the percentage of people holding a minority position

192. See Brian Mullen et al., *Group Cohesiveness and Quality of Decision Making*, 25 *SMALL GROUP RES.* 189, 199-202 (1994); Brian Mullen & Carolyn Copper, *The Relation Between Group Cohesiveness and Performance: An Integration*, 115 *PSYCHOL. BULL.* 210, 225 (1994). For evidence to this effect, see Harrington, *supra* note 35, at 30-34.

193. JAMES S. FISHKIN, *THE VOICE OF THE PEOPLE* 43, 161-81 (1997).

194. *Id.* at 206-07.

195. *Id.*

196. *Id.* at 207.

197. Fishkin & Luskin, *supra* note 6, at 23.

198. See *id.* at 22-23 (showing a jump, on a scale of 1 to 4, from 3.51 to 3.58 in intensity of commitment to reducing the deficit; showing a jump, on a scale of 1 to 3, from 2.71 to 2.85 in intensity of support for greater spending on education; and showing a jump, on a scale of 1 to 3, from 1.95 to 2.16, in commitment to aiding American business interests abroad).

(with, for example, a jump from 36% to 57% of people favoring policies making divorce “harder to get”).¹⁹⁹ These are not the changes that would be predicted by group polarization.

Several factors distinguish the deliberative opinion poll from experiments on group polarization. First, Fishkin’s deliberators did not vote as a group, and while group polarization is observed when no group decision is expected,²⁰⁰ the extent of polarization may well decrease simply because members have not been asked to sign onto a group decision as such. Second, Fishkin’s groups were overseen by a moderator; this attempt to ensure a level of openness is likely to have altered some of the dynamics discussed here. Third, Fishkin’s groups were highly diverse and enclave deliberation was impossible. Fourth, Fishkin’s studies presented participants with a set of written materials that attempted to be balanced and that contained detailed arguments for 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 authoritative external materials. Indeed, the very effort to produce balance should be expected to shift large majorities into small ones, pressing both sides closer to 50% representation; and this is in fact what was observed in many of the outcomes in deliberative opinion polls.²⁰¹

In short, external materials shift the argument pool available to the deliberators and are also likely to have effects on social influence. Once certain arguments are on the table, it is harder to say how one or another position will affect a group member’s reputation. The most sensible conclusion is that externally provided information, the existence of monitors, an absence of a group decision, and the great heterogeneity of the people involved in Fishkin’s studies²⁰² make the deliberative opinion poll quite different from the group polarization studies.

Fishkin’s experiments suggest that group polarization can be heightened, diminished, or possibly even eliminated by seemingly small alterations in institutional arrangements. To the extent that limited argument pools and social influences are likely to have unfortunate effects, correctives can be introduced, perhaps above all by exposing group members, at one point or another, to arguments to which they are not antecedently inclined. Current proposals would do well to incorporate an

199. *Id.* at 23; *see also id.* at 22 (showing an increase, on a scale of 1 to 3, from 1.40 to 1.59 in commitment to spending on foreign aid; also showing a decrease, on a scale of 1 to 3, from 2.38 to 2.27 in commitment to spending on Social Security).

200. *See* Teger & Pruitt, *supra* note 76, at 201-02 (finding group polarization on mere exposure to the views of others).

201. *Id.* at 196-201.

202. Note, however, that choice shifts were produced in the context of jury deliberations, also involving heterogeneous groups. Schkade et al., *supra* note 8, at 1154-56.

understanding of these sometimes neglected facts.²⁰³ But the most important lesson is the most general: It is desirable to create spaces for enclave deliberation without insulating enclave members from those with opposing views, and without insulating those outside of the enclave from the views of those within it.

VII. CONCLUSION

In this Essay, I have discussed the phenomenon of group polarization and explored some of its implications for deliberation generally and deliberative democracy in particular. The central empirical finding is that group discussion is likely to shift judgments toward a more extreme point in the direction indicated by the median of predeliberation judgments. This is true if a group decision is required; if individuals are polled anonymously afterwards, they are likely to shift in precisely the same way.

The underlying mechanisms are twofold. The first involves people's desire to stand in a particular relation to the group, perhaps for reputational reasons, perhaps to maintain their self-conception. Shifts occur as people find that it is necessary to alter their positions in order to maintain their self-conception or their desired relation to the group. The second mechanism involves limited "argument pools," as members of groups with a certain initial tendency typically hear a large number of arguments in support of that tendency, and few arguments in the other direction. When arguments are skewed toward a particular point of view, group members will move in the direction of that point of view. In a finding of special importance to democratic theory, group polarization is heightened if members have a sense of shared identity. And in an equally important finding, group polarization is diminished, and depolarization may result, if members have a degree of flexibility in their views and groups consist of an equal number of people with opposing views.

In the abstract, and without knowing anything about the underlying substance, it is impossible to say whether group polarization is good or bad. But the mechanisms that underlie group polarization raise serious questions about the view that deliberation is likely to yield correct answers to social questions. Like-minded people engaged in discussion with one another may lead each other in the direction of error and falsehood, simply because of the limited argument pool and the operation of social influences. This point very much bears on deliberation within insulated groups and hence on

203. Hence there is reason for caution about the proposal for "deliberation day" offered by Bruce Ackerman and James Fishkin. BRUCE ACKERMAN & JAMES FISHKIN, *DELIBERATION DAY* (forthcoming). The problem lies in the highly territorial nature of the proposal, in which people would deliberate with those in their community. This is a recipe for group polarization and social fragmentation.

emerging communications technologies, which allow a high degree of individual filtering;²⁰⁴ insulation and filtering can compound error. The point also bears on the design of deliberating courts, legislatures, and regulatory agencies. Above all, an understanding of group polarization helps explain why like-minded people, engaged in deliberation with one another, sometimes go to astonishing extremes and commit criminal or even violent acts.²⁰⁵

This is the dark side of “enclave deliberation.” But I have also emphasized that deliberation within protected enclaves can be highly desirable. Partly as a result of group polarization, enclave deliberation can produce positions that would otherwise fail to emerge and that emphatically deserve a public hearing. The case for enclave deliberation is strengthened by evidence that members of low-status groups are likely to be silent in, or silenced by, broader deliberating bodies. Group polarization within enclaves might even operate as a counterweight to this problem.

In the abstract, it is not possible to specify the appropriate mix of enclave deliberation and deliberation within larger publics. But an appreciation of group polarization helps show why a free society takes steps to protect deliberation within enclaves, to ensure that those inside enclaves hear alternative views, and to ensure as well that those outside of particular enclaves are exposed to what enclave members have to say. Above all, it is important to avoid a situation in which people are exposed to softer and louder echoes of their own voices.

In a heterogeneous society, this form of self-insulation can create serious deliberative trouble, in the form of mutual incomprehension or much worse. Legal arrangements will increase or reduce that trouble. I have outlined some approaches that might ensure that heterogeneity, far from being a source of social fragmentation, will operate as a creative force, helping to identify problems and even solutions that might otherwise escape notice.

204. See LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE 186 (2000); SHAPIRO, *supra* note 127, at 105-23.

205. The case of like-minded deliberators is the simplest one, but note that as long as there is a defined median, group polarization can occur in heterogeneous groups as well.