



The Law of Group Polarization

The Harvard community has made this article openly available. [Please share](#) how this access benefits you. Your story matters

Citation	Cass R. Sunstein, The Law of Group Polarization (John M. Olin Program in L. & Econ. Working Paper No. 91, 1999).
Published Version	http://chicagounbound.uchicago.edu/law_and_economics/542/
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:13030952
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

1999

The Law of Group Polarization

Cass R. Sunstein

Follow this and additional works at: http://chicagounbound.uchicago.edu/law_and_economics

 Part of the [Law Commons](#)

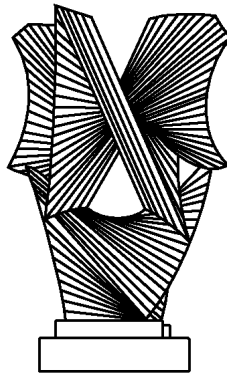
Recommended Citation

Cass R. Sunstein, "The Law of Group Polarization" (John M. Olin Program in Law and Economics Working Paper No. 91, 1999).

This Working Paper is brought to you for free and open access by the Coase-Sandor Institute for Law and Economics at Chicago Unbound. It has been accepted for inclusion in Coase-Sandor Working Paper Series in Law and Economics by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.

CHICAGO

JOHN M. OLIN LAW & ECONOMICS WORKING PAPER NO. 91
(2D SERIES)



The Law of Group Polarization

Cass R. Sunstein

THE LAW SCHOOL
THE UNIVERSITY OF CHICAGO

This paper can be downloaded without charge at:

The Chicago Working Paper Series Index:
<http://www.law.uchicago.edu/Publications/Working/index.html>

The Social Science Research Network Electronic Paper Collection:
http://papers.ssrn.com/paper.taf?abstract_id=199668

The Law of Group Polarization

Cass R. Sunstein*

Abstract

In a striking empirical regularity, deliberation tends to move groups, and the individuals who compose them, toward a more extreme point in the direction indicated by their own predeliberation judgments. For example, people who are opposed to the minimum wage are likely, after talking to each other, to be still more opposed; people who tend to support gun control are likely, after discussion, to support gun control with considerable enthusiasm; people who believe that global warming is a serious problem are likely, after discussion, to insist on severe measures to prevent global warming. This general phenomenon -- group polarization -- has many implications for economic, political, and legal institutions. It helps to explain extremism, "radicalization," cultural shifts, and the behavior of political parties and religious organizations; it is closely connected to current concerns about the consequences of the Internet; it also helps account for feuds, ethnic antagonism, and tribalism. Group polarization bears on the conduct of government institutions, including juries, legislatures, courts, and regulatory commissions. There are interesting relationships between group polarization and social cascades, both informational and reputational. Normative implications are discussed, with special attention to political and legal institutions.

"The differences of opinion, and the jarrings of parties in [the legislative] department of the government . . . often promote deliberation and circumspection; and serve to check the excesses of 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 the standpoint of others 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. The author is grateful to Timur Kuran, Andrei Marmor, Eric Posner, and Richard Posner for valuable comments, and to David Schkade and Daniel Kahneman for many helpful discussions. Participants in a work-in-progress lunch at the University of Chicago also provided a great deal of help.

¹ The Federalist No. 70, at 426-37 (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

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, inclined to be supportive of affirmative action, meet to exchange views and to plan further action, if necessary. 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 local group of citizens, all of them Republicans, meet in 1998 to discuss whether President Clinton should be impeached. Before discussion begins, a strong majority is leaning in favor of impeachment, but they are not firmly committed to this view. A minority is entirely undecided. If a group resolution is required, what is it likely to look like?
- 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, the 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?
- A group of women are 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. After a year, is it possible to say what its members are likely to think?
- There is an Internet discussion group, consisting of people concerned about the behavior of certain activities by Americans apparently associated with

³ Patricia Wallace, *The Psychology of the Internet* 81-82 (1999).

China. Over half of the participants are fearful that China might be engaged in spying and that under President Clinton, the Department of Justice has turned a blind eye, in part because of campaign contributions from Americans whose loyalties are suspect. In what directions are these Internet discussions likely to lead?

Every society contains innumerable deliberating groups. Faculties, juries, legislative bodies, political organizations, regulatory commissions, multimember courts, faculties, student organizations, religious sects, Internet discussion groups, and others engage in deliberation. A pervasive question has to do with the likely consequences of the deliberative process. It is a simple social fact that sometimes people enter discussions with one view and leave with another, even on political and moral 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 much dealt with the real-world consequences of deliberation, and with whether any generalizations hold in actual deliberative settings.

The standard view of deliberation is that of Hamilton and Rawls, as stated 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 groups “all come together . . . they may surpass – collectively and as a body, although not individually – the quality of the few best. . . . When there are many who contribute to the process of deliberation, each can bring his share of goodness and moral prudence; . . . some appreciate one part, some another, and all together appreciate all.”⁷ An important question is whether this view is naïve or excessively optimistic. Perhaps economic, psychological, and social mechanisms lead deliberating groups in unexpected and undesirable directions. If so, it would be necessary to rethink current enthusiasm for deliberation as a social phenomenon, and also to reassess and perhaps to restructure institutions that are designed as deliberating bodies.

My principal purpose in this Article is to investigate a striking but thus far almost entirely neglected⁸ empirical regularity – that of **group polarization** -- and to relate this phenomenon to a number of issues in law and political theory. In brief, group polarization arises when **members of a deliberating group move toward a more extreme point in whatever direction is**

⁴ Sometimes it may seem that moral and political arguments are unlikely to have an effect; the evidence discussed here shows that on this proposition is quite wrong as an empirical matter.

⁵ See Amy Gutmann and Dennis Thompson, *Democracy and Disagreement* (1997); *Deliberative Democracy* (Jon Elster ed. 1998); Jürgen Habermas, *Between Law and Norms* (1997).

⁶ Exceptions include James Fearon, *Deliberation As Discussion*, in *Deliberative Democracy*, supra, at 44; Susan Stokes, *Pathologies of Deliberation*, in id. at 123; Lynn Sanders, *Against Deliberation*, *Political Theory*. Of special interest is 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 Fishkin, *The Voice of the People* (1998); *The Poll With A Human Face* (Maxwell McCombs and 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. In the experiments discussed here, the relevant arguments are introduced by the participants, not by any third party. See below for discussion of Fishkin.

⁷ Aristotle, *Politics* 123 (E. Barker trans. 1972).

⁸ I have been unable to find any sustained discussions in the relevant literature in law or political theory.

indicated by the members' pre-deliberation tendency. “[L]ike polarized molecules, group members become even more aligned in the direction they were already tending.”⁹ Group polarization is the conventional consequence of group deliberation. Thus, for example, the first deliberating group is likely to become more firmly committed to affirmative action; the second group will probably end up favoring gun control quite enthusiastically; any group resolution from the third group will tend to favor impeachment; the punitive damages jury will likely come up with an award higher than the median and perhaps higher than the mean as well; the group of women concerned about feminism is likely to become very conservative indeed on gender issues; the Internet group is likely to fear something like a conspiracy to cover up the relevant activities.¹⁰

Two principal mechanisms underlie group polarization. The first points to social influences on behavior; the second emphasizes limited “argument pools,” and the directions in which those limited pools lead group members. An understanding of these mechanisms provides many insights into legal and political issues; it illuminates a great deal, for example, about likely processes within multimember courts, juries, political parties, and legislatures – not to mention insulated ethnic groups, extremist organizations, student associations, faculties, workplaces, and families. At the same time, these mechanisms give little reason for confidence that deliberation is making things better than worse; in fact they raise some serious questions about deliberation from the normative point of view.¹¹ If deliberation simply pushes a group toward a more extreme point in the direction of its original tendency, do we have any systematic reason to think that discussion is producing improvements?

As we will see, one of the principal lessons of the group polarization phenomenon is to cast new light on an old point, to the effect that social homogeneity can be quite damaging to good deliberation.¹² When people are hearing echoes of their own voices, the consequence may be far more than support and reinforcement. Another lesson is that particular forms of homogeneity can be breeding grounds for unjustified extremism, even fanaticism. To work well, deliberating groups should be appropriately heterogeneous and should contain a plurality of articulate people with reasonable views – an observation with implications for the design of regulatory commissions, legislative committees, White House working groups, and even multimember courts.¹³ But there is a conceptual problem here: It is difficult to specify appropriate heterogeneity, and the appropriate plurality of views, without making some antecedent judgments about the substantive question at issue. I offer some comments about how to resolve that problem.

This Article is organized as follows. Part II offers some brief notations on the general question of social influences on individual judgments, with particular reference to the

⁹ See John Turner et al., *Rediscovering the Social Group* 142 (1987).

¹⁰ Compare R. Hightower and L. Sayeed, *The Impact of Computer-Mediated Communication Systems on Biased Group Discussion*, 11 *Computers in Human Behavior* 33 (1995).

¹¹ I am speaking here of real-world deliberation, not of deliberation accompanied by preconditions of the sort that have been influenced by those thinking of it in ideal terms. See Jurgen Habermas, *supra*. A particular point to emphasize here is the need for full information. See *id.*

¹² A classic discussion is John Stuart Mill, *On Liberty* (1859).

¹³ Compare Irving Janis, *Groupthink* (1972) (coming to the same general conclusion, but without discussing polarization).

phenomenon of social cascades. A central point here is that any particular's persons deeds and statements create an informational externality.¹⁴ When a number of people have acted or spoken, observers who lack much private information are highly likely to follow their lead. Part III offers a basic account of group polarization, with particular reference to some new data in the legal context. Part IV discusses the mechanisms that account for group polarization. Part V traces the implications for a number of issues, involving feuds, ethnic strife, juries, commissions, multimember courts, legislatures, and deliberation via the Internet. Part VI shows in what sense group polarization raises doubts about the idea that deliberation is a social good; it traces the implications of the phenomenon for proper structuring of deliberative institutions. 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.¹⁵ Because many of these influences are at least roughly analogous to what happens in group polarization, and because they have some bearing on deliberation as well, it will be worthwhile to offer some brief notations here.

The simplest point is that people frequently do what they do because of what they think (relevant) others do. Thus, for example, teenage girls who see that other teenagers are having babies are more likely to become pregnant themselves¹⁶; littering and nonlittering behavior appears to be contagious¹⁷; the same is true of violent crime¹⁸; those who know other people who are on welfare are more likely to go on welfare themselves¹⁹; the behavior of proximate others affects the decision whether to recycle²⁰; a good way to increase the incidence of tax compliance is to inform people of high levels of voluntary tax compliance²¹; 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 this fact is one of the few successful methods of reducing binge drinking on college campuses.²²

¹⁴ See Andrew Caplin & John Leahy, *Miracle on Sixth Avenue: Information Externalities and Search*, 108 *Econ. J.* 60 (1998).

¹⁵ See, e.g., Dan Kahan, *Social Influence, Social Meaning, and Deterrence*, *Va L. Rev.* (1998). For extended overviews, see Elliott Aronson, *The Social Animal* (7th ed. 1995); Lee Ross and Richard Nisbett, *The Person and the Situation* (1991); group polarization is a surprising omission from both of these lengthy and highly illuminating treatments.

¹⁶ See, e.g., George A. Akerlof, Janet L. Yellen & Michael L. Katz, *An Analysis of Out-of-Wedlock Childbearing in the United States*, 111 *Q.J. Econ.* 277 (1996).

¹⁷ See Robert Cialdini et al., *A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places*, 58 *J. Pers. And Soc. Psych.* 1015 (1990).

¹⁸ See *Washington Post* (December 1999).

¹⁹ See Marianne Bertrand, Erzo F.P. Luttmer & Sendhil Millainathan, *Network Effects and Welfare Cultures* (unpublished manuscript, Apr. 9, 1998).

²⁰ See Ardith Spence, *Wants for Waste* (unpublished Ph.D. dissertation, University of Chicago, 1999).

²¹ See Stephen Coleman, *The Minnesota Income Tax Compliance Experiment State Tax Results* (Minnesota Department of Revenue, April 1996).

²² See H. Wesley Perkins, *College Student Misperceptions of Alcohol and Other Drug Norms Among Peers*, in *Designing Alcohol and Other Drug Prevention Programs in Higher Education* 177-206 (US Dept. of Educ. ed. 1997); Timur Kuran and Cass R. Sunstein, *Availability Cascades and Risk Regulation*, 51 *Stan L. Rev.* 683, 767 (1999).

Social influences affect behavior via two different mechanisms.²³ The first is informational. As noted, what other people do, or say, carries an informational externality; if many other people go to a certain movie, or refuse to use drugs, or carry guns, observers are given a signal about what it makes sense to do. The second mechanism is reputational. Even if people do not believe that what other people do provides information about what 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 hence they may do what they think other people think they should do, whether or not they believe that they should do it. Reputational considerations may, for example, lead people to obey or not to obey the law, smoke cigarettes, buy certain cars,²⁴ drive while drunk, help others, or talk about political issues in a certain way. They exert a ubiquitous influence on behavior.²⁵

B. Some Classic Experiments

In the most vivid experiments involving group influences, conducted by Solomon Asch, individuals were willing to abandon the direct evidence of their own senses.²⁶ 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, 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 experiments was one of eight people asked to engage in the matching. But unbeknownst to the subject, the other people apparently being tested were actually there as part of the experiments.

Asch's experiments unfolded in the following way. In the first two rounds, everyone agreed about the right answer; this seemed to be an extremely dull experiment. But the third round introduced "an unexpected disturbance,"²⁷ 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 percent of the time; but in rounds in which group pressure supported the incorrect answer, subjects erred 36.8% of the time.²⁸ Indeed, in a series of twelve questions, no less than 70% of subjects went along with the group, and defied the evidence of their own senses, at least once.

²³ See, e.g., Elliott Aronson, *supra* note, at 22; Lee and Ross, *supra* note, at 44-45.

²⁴ See Robert Frank, *Luxury Fever* (1999).

²⁵ Timur Kuran, *Public Truth, Private Lies* (1998), emphasizes this point.

²⁶ See the overview in Solomon Asch, *Opinions and Social Pressure*, in *Readings About the Social Animal* 13 (Elliott Aronson ed. 1995).

²⁷ *Id.* at 15.

²⁸ *Id.* at 16.

Several refinements are important here. 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, the existence of at least one compatriot, or voice of sanity, mattered a great deal. 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.²⁹ By contrast, varying the size of the majority mattered only up to a number of three, and increases from that point had little effect. Thus opposition from one person did not increase subjects' errors at all; opposition from two people increased error to 13.6%; and opposition from three people increased error to 31.8%, not substantially different from the level that emerged from further increases in group size.

Both informational and reputational considerations appear to have led people toward these errors. Several people said, in private interviews, that their own opinions must have been wrong. On the other hand, experimenters find greatly reduced error, in the same basic circumstances as Asch's experiments, when the subject is asked to give a purely private answer.³⁰

Asch concluded that his results raised serious questions about the possibility that "the social process is polluted" by the "dominance of conformity."³¹ He added, "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."³² Notably, however, Asch's experiments did not involve deliberation, for people were not exchanging reasons; indeed, we might expect that reason-giving would have severely weakened his results. 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 will lead people in the right directions.

C. Cascades

Some of the most interesting recent work on social influence involves the possibility of informational and reputational "cascades"³³; this work has obvious relevance to law and politics.³⁴ Indeed, it is possible to interpret Asch's work as having demonstrated considerable individual susceptibility to cascade effects. What is striking about such effects is that their ripple-like nature, or the quality of contagion. Group polarization is sometimes, but not always, a product of cascade effects; it will be useful to understand the former against the background of the latter.

The question explored in the cascades literature is why individuals and social groups sometimes move quite rapidly in some direction or another. A starting point is that when individuals lack a great deal of private information (and sometimes even when they have

²⁹ Id. at 18.

³⁰ See Aronson, *supra* note, at 23-24.

³¹ Id. at 21.

³² Id.

³³ See Sushil Bikhchandani et al., *Learning from the Behavior of Others*, *J. Econ. Persp.*, Summer 1998, at 151

³⁴ See *id.*; Timur Kuran and Cass R. Sunstein, *Availability Cascades and Risk Regulation*, 51 *Stan L Rev* (1999).

such information), they tend to 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 believe that fear is justified, C may end up thinking so too, at least if she lacks independent information to the contrary. If A, B, and C believe that abandoned hazardous waste dumps are hazardous, D will have to have a good deal of confidence to reject their shared conclusion. The result of this process can be to produce cascade effects, as large groups of people end up believing something – even if that something is false – simply because other people seem to believe it too. There is a great deal of experimental evidence of informational cascades, which are easy to induce in the laboratory³⁵; real world phenomena also seem to have a great deal to do with cascade effects.³⁶ Notice here that when a cascade is occurring, large numbers of persons end up with a shared view, not simply because of social influence, but via a particular process, in which a rivulet ends up as a flood; this is what makes cascades distinctive.

Though the cascades phenomenon has largely been discussed in connection with factual judgments, the same processes should be at work for political, legal, and moral questions; we can easily imagine political, legal, and moral cascades. Suppose, for example, that A believes that affirmative action is wrong, that B is otherwise in equipoise but shifts upon hearing what A believes, that C is unwilling to persist in his modest approval of affirmative action when A and B disagree; it would be a very confident D who would reject the moral 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, may end up believing what (relevant) others seem to believe. Recent changes in social attitudes toward smoking, recycling, and sexual harassment have a great deal to do with these effects.³⁷ The same process may 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. We can easily imagine cascade effects in the direction of certain judgments about the appropriate course of 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 (eg, textualism, originalism).

Thus far the discussion has involved purely informational pressures and informational cascades, where people care about what other people think because they do not know what to think, and they rely on the opinions of others, to show what it is right to think. But there can be reputational pressures and reputational cascades as well.³⁸ Here the basic idea is that people care about their reputations, and they speak out, or remain silent, or even engage in certain expressive activity, partly in order to preserve those reputations, even at the price of failing to say what they really think. Suppose, for example, that A believes that hazardous waste dumps pose a serious

³⁵ See Lisa Anderson and Charles Holt, *Information Cascades in the Laboratory*, 87 *Am Econ Rev* 847 (1997).

³⁶ See Bikhchandani et al., *A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades*, 100 *J Polit Econ*. 992 (1992); Kuran and Sunstein, *supra* note.

³⁷ See Spence, *supra* note; Cass R. Sunstein, *Social Norms and Social Roles*, in Cass R. Sunstein, *Free Markets and Social Justice* ch. 2 (1997).

³⁸ See Timur Kuran, *Public Lies, Private Truths* (1996).

environmental problem; suppose too that B is skeptical. B may keep quiet, or (like some of Asch's subjects) even agree with A, simply in order to preserve A's good opinion. C may see that A believes that hazardous waste dumps pose a serious problem, and that B seems to agree with A; C may therefore voice agreement even though privately she is skeptical or ambivalent. It is easy to see how this kind of thing might happen with in political life with, for example, politicians expressing their commitment to capital punishment (even if they are privately skeptical) or their belief in God (even if they are agnostic on the question). Here too the consequence can be cascade effects – large social movements in one direction or another -- when a number of people appear to support a certain course of action simply because others (appear to) do so. What is true for factual beliefs can 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.

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 public torpor, by generating concern about serious problems; but sometimes cascade effects will make people far more worried than they be, or otherwise produce large-scale distortions in private judgments, public policy, and law. The antislavery movement had distinctive cascade-like features, as did the anti-apartheid movement in South Africa; so too with Mao's Cultural Revolution and the rise of Nazism in Germany.³⁹ 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 we will now see.

III. How and Why Groups Polarize

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 [by] group discussion.”⁴⁰ The result is that groups often make more extreme decisions than would the typical or average individual in the group (where “extreme” is defined internally, by reference to the group's initial dispositions). There is a clear relationship between group polarization and cascade effects; as we will see, the former, like the latter, seems to have a great deal to do with both informational and reputational influences. A key difference is that cascade effects lead people to fall in line with an existing tendency, whereas polarization leads them to a more extreme point in the same direction.

³⁹ See *The Social Life of Nazi Germany* (1999).

⁴⁰ See Isenberg, *supra* note, at 1141.

Notice that group polarization refers not to variance among groups of any kind, but to what happens within a group discussing a case or problem.⁴¹ Consider some examples of the basic phenomenon, which has been found in an array of nations.⁴² (a) A group of moderately profeminist women will become more strongly profeminist after discussion.⁴³ (b) After discussion, citizens of France become more critical of the United States and its intentions with respect to economic aid.⁴⁴ (c) After discussion, whites predisposed to show racial prejudice offer more negative responses to the question whether white racism is responsible for conditions faced by African-Americans in American cities.⁴⁵ (d) After discussion, whites predisposed not to show racial prejudice offer more positive responses to the same question.⁴⁶ As statistical regularities, it should follow, for example, that those moderately critical of an ongoing war effort will, after discussion, sharply oppose the war; that a group moderately predisposed to hire a certain job candidate will, after discussion, support the application with considerable enthusiasm; that people tending to believe in the inferiority of a certain racial group will be entrenched in this belief as a result of discussion.

The phenomenon of group polarization has conspicuous importance to the operation of deliberating bodies of relevance to law and politics, including legislatures, commissions, multimember courts, and juries. I will return to this point shortly; for now notice a few obvious possibilities. Members of a political party, or of the principal political parties, may polarize as a result of internal discussions; party-line voting is sometimes explicable partly on this ground. A set of judges with similar predilections on a three-judge panel may well produce a more extreme ruling than any individual member would write if he were judging on his own. Extremist groups will often become more extreme; as we will soon see, the largest group polarization typically occurs with individuals already inclined toward extremes. With respect to deliberating juries, a recent study⁴⁷ found significant group polarization with respect to “numerical punishment ratings” on a bounded numerical scale. For high punishment ratings, groups tended to generate numbers higher than the median of individual predeliberation judgments; for low punishment ratings, groups tended to generate numbers lower than the median of individual predeliberation judgments. This is precisely the pattern that group polarization would predict.

B. Risky Shifts and Cautious Shifts

Group polarization was first found in a series of experiments involving risk-taking decisions.⁴⁸ Before 1961, conventional wisdom had been that as compared with the individuals who compose it, a group of decision-makers – for example a committee or board – would be

⁴¹ Of course, when different deliberating groups polarize in different directions, the consequence can be great among-group variance.

⁴² These include the United States, Canada, Germany, and France. 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.

⁴³ See D.G. Myers, Discussion-Induced Attitude Polarization, 28 Human Relations 699 (1975).

⁴⁴ Brown, Social Psychology 224 (2d ed. 1983).

⁴⁵ D.G. Myers and G.D. Bishop, The Enhancement of Dominant Attitudes in Group Discussion, 20 J Personality and Soc. Psych. 286 (1976),

⁴⁶ See *id.*

⁴⁷ See David Schakde et al., Are Juries More Erratic Than Individuals?, Colum L Rev (2000).

⁴⁸ I draw in this and the following paragraph on Brown, *supra* note, at 200-206.

likely to favor a compromise and thus to avoid risks. But the relevant experiments, originally conducted by Stoner, found otherwise; they identified what has become known as the “risky shift.”⁴⁹ Deliberation tended to shift group members in the direction of greater risk-taking; and deliberating groups, asked to reach a unanimous decision, were generally more risk-inclined – sometimes far more risk-inclined – than the mean individual member, predeliberation.

In the original experiments, male graduate students of industrial management were asked a range of questions involving risk: whether someone should choose a safe or risky play in the last seconds of a football game; whether someone should invest money in a low-return, high-security stock or instead a high-return, lower security stock; whether someone should choose a high prestige graduate program in which a number of people fail to graduate or a lower prestige school where everyone graduates. In one problem, for example, people were asked to say whether a person now having a secure, lifetime job should take a new job, with a new company with an uncertain future. People were asked about the lowest probability of “financial soundness” that would justify the person with the secure job from taking the new position. In Stoner’s studies, people first studied the problems – twelve total – and recorded an initial judgment; they were then asked to reach a unanimous decision as a group. People were finally asked to state their private judgments after the group judgment had been made; they were informed that it was acceptable for the private judgment to differ from the group judgment.

For twelve of the thirteen groups, the group decisions showed a repeated pattern toward greater risk-taking -- that is, after discussion, the unanimous outcome tended to assess the necessary likelihood of financial soundness as consistently lower than the median judgment of the group predeliberation. In addition, there was a clear shift toward greater risk-taking in private opinions as well. Only 16% were moved toward greater caution; 45% did not change at all; and a full 39% moved in the direction of greater risk-taking. This shift – the “risky” shift – was promptly duplicated in a number of diverse studies, some involving all men and some involving all women.

We should distinguish at this point between two aspects of these findings, not always separated in the psychological literature and both of relevance to law and policy. The first involves the movement of deliberating groups, for whom a group decision is necessary, toward the group’s extreme end; call this (inelegantly) group polarization toward within-group extremes. 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 may well, for example, 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 individual polarization toward within-group extremes. This means that to the extent that private judgments are moved by discussion, it will be toward a more extreme point in the direction set by the original distribution of views.

⁴⁹ See J.A.F. Stoner, A Comparison of Individual and Group Decisions Including Risk, unpublished master’s thesis, School of Management, Massachusetts Institute of Technology; J.A.F. Stoner, Risky and Cautious Shifts in Group Decisions, 4 J Experimental Social Psych. 442 (1968).

A possible (and contemporaneous) reading of Stoner's early studies would be that group dynamics are such as to move people – both groups and individuals within them -- in the direction of greater risk-taking. But this conclusion would be much too simple. Later studies showed that under certain conditions, it was possible, even easy to induce a “cautious shift” as well. Indeed, certain problems reliably produced cautious shifts. The principal examples involved 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 moved toward caution, as did the members who composed them.

As yet there is no simple account of what kinds of problems will produce what kinds of shifts; but the identification of risky and cautious shifts has helped produce a general account of how much, and in what direction, people will tend to move. In Stoner's original data, subsequent researchers noticed, the largest risky shifts could be found when group members “had a quite extreme risky initial 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.”⁵⁰ 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. A group of very cautious individuals would produce a significant shift toward greater caution; a group of individuals inclined toward risk-taking would produce a significant shift toward greater risk-taking; and groups of individuals in the middle would produce smaller shifts in the direction indicated by their original disposition. In short, “group discussion moves decisions to more extreme points in the direction of the original inclination . . . , which means shift to either risk or caution in the direction of the original disposition, and the size of the shift increases with the degree of the initial polarization.”⁵¹ Similar results have been found in many contexts, involving, for example, questions about economic aid, architecture, political leaders, race, feminism, and judgments of guilt or innocence.⁵² Polarization has been found for questions of obscure fact (eg, how far Sodom on the Dead Sea is below sea level) as well as for evaluative questions, including political and legal issues⁵³ and even the attractiveness of people in slides.⁵⁴

IV. Mechanisms

A. Two Mechanisms

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

⁵⁰ Brown at 211.

⁵¹ Brown, *supra*, at 211.

⁵² See *id.*

⁵³ A relatively recent treatment is Russell Spears, Martin Lee, and Stephen Lee, De-Individuation and Group Polarization in Computer-Mediated Communication, 29 *British J Soc Psych* 121 (1990)

⁵⁴ Turner et al., *supra*, at 153.

There have been two main explanations for group polarization, both of which have been extensively investigated.⁵⁵ Massive support has been found on behalf of both explanations.⁵⁶

1. Social comparison. The first, involving social comparison, begins with the claim that people want to be perceived favorably by other group members, and also to perceive themselves favorably. Once they hear what others believe, they adjust their positions in the direction of the dominant position. They may want to signal, for example, that they are not cowardly or cautious, and hence they will frame their position so that they do not appear such by comparison to other group members.⁵⁷ With respect to risk-taking activity, people want to occupy a certain position in comparison to others, and before they hear what other people think, they assume that they do in fact occupy that position. But when they hear what other people think, they find, often, that they occupy a somewhat different position, and they shift accordingly. The result is to press the group's position toward one or another extreme, and also to induce shifts in individual members. The same appears to happen in other contexts. People may wish, for example, not to seem too enthusiastic, or too restrained in their enthusiasm for, affirmative action, feminism, or an increase in national defense; hence their views may shift when they see what other group members think. The result will be both group and individual polarization toward within-group extremes.

The dynamic behind the social comparison explanation is that most people may want 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 moderate risk-takers, and their choice of position is partly a product of this desire.⁵⁸ No one can know what such a position would be until the positions of others are revealed.⁵⁹ Thus individuals move their judgments in order to preserve their image to others and their image to themselves. A key claim here is that information alone about the actual positions of others – without discussion -- will produce a shift. Evidence has confirmed this fact; mere exposure induces a substantial risky shift (though it is less substantial than what is produced by discussion – about half as large).⁶⁰ This effect helps explain a shift toward caution (the “cautious shift”) as well.⁶¹ While highly suggestive, the “mere

⁵⁵ Isenberg, *supra*, and Brown, *supra*, review this literature; see also Turner et al., *supra*, at 142-70, for an overview and an attempt to generate a new synthesis.

⁵⁶ Note that conformity does not explain group polarization. People are not attempting to conform, even under the social comparison theory; they are attempting to maintain their relative position, and the revelation of the views of others shifts people's conception of what judgment is necessary to maintain that position. See Myers, *supra* note, at 562, indicating that people “want to perceive themselves as somewhat different from others” and that “people want to differentiate themselves from others, to a small extent and in the right direction.”

⁵⁷ On signalling generally, see Eric Posner, *Symbols, Signals, and the Law* (forthcoming 2000).

⁵⁸ For a quite vivid demonstration of such a process in the enactment of the Clean Air Act, one that does not, however, identify the mechanisms, discussed here, see Bruce Ackerman, John Millian, and Donald Elliott, *Toward a Theory of Statutory Evolution: The Federalization of Environmental Law*, 1 *J. L. Econ. & Organization* 313 (1985).

⁵⁹ “Once the real locations 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*, at 214.

⁶⁰ Teger and Pruitt (1967).

⁶¹ Investigations of social influence have emphasized both one-upmanship and the removal of pluralistic ignorance, that is, ignorance of what other people think (or are willing to say they think). Note that it is implicit in these findings that people seem to want not to conform, but to be different from others in a desirable way. “To be virtuous . . . is to be different from the mean – in the right direction and to the right degree.” Brown, *supra* note, at 469.

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. (Recall the discussion of informational cascades.)

The social influence explanation invokes factors similar to those that underlie the reputational cascade. A major difference is that the social influence explanation concerns presentation to self as well as presentation to others. Note also that group polarization may or may not be a result of any cascade effect; the question is whether the accumulation of views from others operates in the form of a cascade. Existing work on group polarization does not answer this question.

2. Persuasive arguments. The second explanation, emphasizing the role of persuasive arguments, is based on a common sense 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 collectivity. Because a group whose members are already inclined in a certain direction will have a disproportionate number of arguments supporting that same direction, the result of discussion will be to move individuals 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.

The persuasive arguments theory begins with the suggestion that if a group is deliberating about some difficult question with a factual answer (how many countries are there in Africa, for example, or how many people were on the planet in 1900), discussion will typically produce some movement, not toward the mean, but toward the minority view on which one or a few members have accurate information. There is, moreover, empirical evidence that with respect to facts, deliberation produces movements toward accuracy.⁶² Of course many of the questions involving group polarization do not have purely factual answers. But a key aspect of those discussions is that the person with the correct answer is likely to state his view with a high degree of confidence, and also be able to make some argument in favor of that view. Novel arguments, bringing up fresh points, are especially likely to be persuasive. In any case members of a group will have thought of some, but not all, of the arguments that justify their initial inclination; consider the question whether to take risks or to be cautious. In discussion, arguments of a large number of individuals are stated and heard, but the total argument pool will be tilted in one or another direction, depending on the predispositions of the people who compose the group; hence there will be a shift in the direction of the original tilt.⁶³

When people hear arguments that they perceive as valid, or find to be memorable, vivid, new, or weighty simply by virtue of emphasis and repetition, they will shift in the direction suggested by those arguments. If a group of moderately feminist women becomes more feminist, a group moderately opposed to gun control more extremely so, and so forth, one reason is that the argument pool of any such group will contain a preponderance of arguments in the direction

⁶² See James Fishkin and Robert Luskin, Bringing Deliberation to the Democratic Dialogue, in *The Poll With A Human Face* 3, 29-31 (Maxwell McCombs and Amy Reynolds eds. 1999).

⁶³ Brown, *supra*, at 219.

suggested. The suggestion is that group polarization will occur when convincing arguments produce a shift in the direction of prediscussion inclinations, revealed in the means of the initial decisions

There is an obvious analogy here to the informational cascade. In fact we can safely assume that group polarization sometimes occurs via a kind of informational cascade, as the statements of particular people begin a cascade process that culminates in extremism. The difference is that for cascade effects, what is crucial is the very fact of the belief, not its grounds, whereas for persuasive arguments to work, what is crucial is that arguments be offered and be found persuasive. It is also unclear whether any particular group polarization involves cascade effects at all; undoubtedly what sometimes happens is not a cascade effect, in which a large number of people successively “fall,” but a simple accumulation of arguments, eventually imposing weight on people whose views are subject to change.

B. Refinements -- and Depolarization

These are statistical regularities, no more. Of course not all groups polarize; some groups end up in the middle, not toward either extreme. Note that in Stoner’s original experiments, one of the twelve deliberating groups showed no polarization at all. Nor is it hard to understand why this might be so. If the people defending the original tendency are particularly unpersuasive, group polarization is unlikely to 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.⁶⁴ In addition, affective factors appear to be quite important and complementary to persuasive arguments. People are less likely to shift if the direction advocated is being pushed by unfriendly group members; the chance of shift is increased when people perceive fellow members as friendly, likeable, and similar to them.⁶⁵ Physical spacing tends to reduce polarization; a sense of common fate and intragroup similarity tend to increase it, as does the introduction of a rival “outgroup.”⁶⁶ Part of the reason for group polarization appears to be that as a class, extreme positions tend to be less tractable and more confidently held. This point is an important complement to the persuasive arguments theory⁶⁷: The persuasiveness of arguments depends, not surprisingly, not simply on the grounds given, but also on the confidence with which they are

⁶⁴ This is of course the theme of the movie *Twelve Angry Men*, where the single hold-out, played by Henry Fonda, shifts the judgment of the jury.

⁶⁵ See Hans Brandstatter, Social Emotions in Discussion Groups, in *Dynamics of Group Decisions* (Hans Brandstatter et al. ed. 1978). Turner et al., *supra* note, at 154-59, 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. The key point here is that when a group is tending in a certain direction, the perceived “prototype” is determined by where the group is leaning, and this is where individuals will shift. *Id.* at 156. As Turner et al. suggest, their account shows “overlap with many aspects of social comparison and persuasive arguments models,” and because of the overlap, I do not discuss it as a separate account here. 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. See *id.* at 158.

⁶⁶ See Turner et al., *supra* note, at 153.

⁶⁷ See Maryla Zaleska, The Stability of Extreme and Moderate Responses in Different Situations, in *Group Decision Making*, *supra*, at 163, 164.

articulated. (Consider here both juries and multimember courts.) 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 can be greatly aggravated.

Notably, the persuasive arguments theory implies that there will be “depolarization,” or convergence toward the middle, if and when new persuasive arguments are offered that are opposite to the direction initially favored by group members. There is evidence for this phenomenon as well.⁶⁸ Depolarization, rather than polarization, will also be found when the relevant group consists of individuals drawn equally from two extremes (a point to which I will return).⁶⁹ Thus 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 positions generally change little as a result of discussion or shift to a more moderate position.⁷⁰ Consider a study⁷¹ consisting of six-member groups specifically designed to contain two subgroups (of three persons each) initially committed to opposed extremes; the effect of discussion was to produce movement toward the center. One reason may be the existence of partially shared persuasive arguments in both directions.⁷² Interestingly, this study of opposed subgroups found the greatest depolarization with obscure matters of fact (e.g., the population of the United States in 1900) -- and the least depolarization with highly visible public questions (e.g., whether capital punishment is justified). Matters of personal taste depolarized a moderate amount (e.g., preference for basketball or football, or for colors for painting a room).⁷³

These findings fit well with the persuasive arguments account of polarization. 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 an additional discussion is not likely to produce movement. Hence “familiar and long-debated issues do not depolarize easily.”⁷⁴ With respect to such issues, people are simply less likely to shift at all.

It also matters whether people think of themselves, antecedently or otherwise, as part of a group, with a degree of solidarity. If they think of themselves in this way, group polarization is all the more likely, and it is likely too to be more extreme.⁷⁵ Thus when people are “de-

⁶⁸ A third possibility is that hearing other similar opinions produces greater confidence in individual positions, opening members to a more extreme judgment in the same direction. raised recently by Heath and Gonzales . See Chip Heath and Richard Gonzales, Interaction With Others Increases Decision Confidence But Not Decision Quality: Evidence Against Information Collection Views Of Interactive Decision Making, 61 *Organizational Behavior and Human Decision Processes* 305-326 (1997).

⁶⁹ See H. Burnstein, Persuasion As Argument Processing, in *Group Decision Making* (H. Brandstetter, J.H. Davis, and G. Stocker-Kreichgauer eds., 1982).

⁷⁰ Ferguson and Vidmar, Effects of Group Discussion on Estimates of Risk Levels, 20 *J Pers and Social Psych* 436 (1971).

⁷¹ Brown, *supra*, at 225.

⁷² Amiram Vinokur and Eugene Burnstein, The Effects of Partially Shared Persuasive Arguments on Group-Induced Shifts, 29 *J Personality & Social Psych.* 305 (1974).

⁷³ Amiram Vinokur and Eugene Burnstein, *supra*, at 884.

⁷⁴ Brown, *supra*, at 226.

⁷⁵ See Russell Spears, Martin Lee, and Stephen Lee, De-Individuation and Group Polarization in Computer-Mediated Communication, 29 *British J Soc Psych* 121 (1990); Patricia Wallace, *The Psychology of the Net* (1999).

individualized,” in the sense that the context emphasizes each person’s membership in the social group engaging in deliberation, polarization increases.⁷⁶

An especially interesting experiment attempted to investigate this point by manipulating two variables.⁷⁷ First, some subjects were “de-individualized” by having to work on computers in separate rooms, whereas others were asked to work in a single office with desks facing each others (the “individualized” condition). In the de-individualized condition, visual anonymity was increased. Second, some subjects were given instructions in which group membership was made salient (the “group immersion” condition), whereas others were not (the “individual” condition). For example, subjects in the group immersion conditions were told that their group consisted solely of first-year psychology students, and that they were being tested as group members rather than as individuals. All conditions were held constant in one respect: Every subject was told that people like them tended to support one or another view. The relevant issues involved affirmative action, government subsidies for the theatre, privatization of nationalized industries, and phasing out nuclear power plans.

The results were quite striking. There was the least group polarization in the de-individuated-individual condition; group polarization was greatest in the de-individuated/group immersion condition, when group members met relatively anonymously and when group identity was emphasized. Interestingly, there was no significant difference in polarization between the two individuated conditions (with and without emphasis on group immersion). From this experiment, it is reasonable to speculate that polarization is most likely to occur, and to be most extreme, under circumstances in which group membership is made salient and people have a high degree of anonymity. There is obviously a potential lesson here about the effects of group deliberation on the Internet,⁷⁸ a point to which I will return.

These remarks suggest some general, common-sensical conclusions about how and when group discussion will move predeliberation 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 more members knows the truth; where views are not firmly held, but where there is an initial predisposition, group polarization is the general rule. Undoubtedly generalizations of this sort bear on shifts in individual views among many deliberating bodies.

B. Iterated “Polarization Games”?

The logic of group polarization suggests 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. Thus, for example, if a group of people 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

⁷⁶ Spears et al., *supra*, at 122-124.

⁷⁷ *Id.*

⁷⁸ See Patricia Wallace, *The Psychology of the Internet* 73-79 (1999).

situation in which individuals hold positions more extreme than those of any individual member before the series of deliberations began.

This is only a thought experiment; there appears to be no study of such iterated polarization games. But the hypothesized outcome is less fanciful than it 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, pre-deliberation.⁷⁹ And it is not difficult to think of real-world groups in which the consequence of deliberation, over time, appears to be to shift both groups and individuals to positions that early on, they could not possibly have accepted.⁸⁰ Iterated polarization games seem to be an important real-world phenomenon. But this raises two questions: Why and when do groups stop polarizing? Why and when do they end up at a certain point, or even shift in the opposite direction? Nothing in the literature on group polarization adequately answers these questions.⁸¹ 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, or when the simple self-interest of political leaders produces a shift in direction,⁸² or when new circumstances, of 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 information⁸³; the same processes seem to terminate or to reverse group polarization.

C. A Wrinkle: “Rhetorical Asymmetry”

Interestingly – and in a noteworthy qualification of the general literature on group polarization – the previously discussed study of punitive damage awards by juries found a striking pattern for dollar awards.⁸⁴ For any dollar award above zero, the general effect of deliberation was to increase awards above those of the median voter. Dollar awards did not simply polarize; while higher awards increased dramatically, as compared to the median of predeliberation votes, low awards increased as well. Why is this?

Both the original experiment and a follow-up experiment suggest that this result is a product of a “rhetorical asymmetry” that favors, other things being equal and in any contest, the person or persons urging higher awards. Thus the fact of systematic increases in dollar awards in strongly suggested of a general tendency toward upward movement; a subsequent experiment, limited to University of Chicago law students, confirmed this effect, with a substantial majority of subjects agreeing that it was easier, other things being equal, to argue for higher awards than

⁷⁹ David Schkade, Cass R. Sunstein, and Daniel Kahneman, *Are Juries More Erratic Than Individuals?*, 100 *Colum L Rev* (forthcoming 2000).

⁸⁰ For a relevant discussion of deliberating groups in the 1960s, see James Miller, *Democracy Is In The Streets* (1993); see also Timur Kuran, *Ethnic Norms and Their Transformation Through Reputational Cascades*, 27 *J Legal Stud* 623, 648 (1998).

⁸¹ See Turner et al., *supra*, at 152, suggesting that there is no clear answer to the question for “what range of situations” polarization is predicted.

⁸² See the discussion of legislators below; and note, as inevitably noisy real-world examples, the centrist shift of the Democratic Party in 1991-1992, and the apparent centrist shift of the Republican Party in 1999-2000.

⁸³ See David Hirschleifer, *The Blind Leading The Blind*, in *The New Economics Of Human Behavior* (1997).

⁸⁴ See David Schkade, Cass R. Sunstein, and Daniel Kahneman, *Are Juries More Erratic Than Individuals?*, 100 *Colum L Rev* (forthcoming 2000).

for lower ones.⁸⁵ In our culture, and in light of existing norms, the person favoring the higher amount for punitive damages appears likely to be more convincing than the person favoring the lower amount. It is important to emphasize that this asymmetry operates independently of any facts about the individual case. The reason appears to be that with respect to dollar awards involving a corporate defendant, stronger 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.

Undoubtedly there are many other contexts containing rhetorical asymmetry, and undoubtedly the asymmetry can affect outcomes, as it did in the jury study. A great deal of empirical work remains to be done on this question; too little is known to say why and when such an asymmetry is at work. Existing cultural norms are the underlying source of any rhetorical asymmetry, and such norms vary over space and time, producing shifts in rhetorical asymmetry. In any case it is not difficult to generate possible examples. Legislative judgments about criminal punishment may, for example, involve an asymmetry of exactly this kind. 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. Much remains to be explored. For present purposes the point is that group polarization may be aggravated or attenuated if one or another side has a systematic advantage in rhetoric. Perhaps the most striking implication is that when there is an initial distribution of views in a certain direction, and when a more extreme movement in that direction has a rhetorical advantage, quite extreme shifts can be expected.⁸⁶

D. Is Group Polarization Rational?

In both economics and law, a great deal of attention has recently been paid to the question whether human beings are “rational,” or “quasi-rational,” or subject to irrationality.⁸⁷ There is an obvious question whether the phenomenon of group polarization raises doubts about rational actor models in economics or law. The answer is that for the most part, individual behavior within groups, as described thus far, creates 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.⁸⁸ More difficult questions might seem to be raised by “social influence” accounts of group polarization. But it is certainly rational for people to care about their reputations. If they are changing their assessment because of reputational considerations, what must be said is that maintaining a certain reputation is 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

⁸⁵ See *id.* at XX.

⁸⁶ Data from *id.* strongly support this speculation, with many deliberating juries producing dollar awards higher, and sometimes significantly higher, than that of the higher individual pre-deliberation. Data to be added.

⁸⁷ Within economics, see, e.g., Gary Becker, *Accounting For Tastes* (1998); Richard Thaler, *Quasi-Rational Economics* (1993). Within law, see Symposium, *The Legal Implications of Psychology*, 51 *Vand. L. Rev.* 1499 (1998); *Behavioral Law and Economics* (Cass R. Sunstein ed. 2000).

⁸⁸ Compare Hirshleifer, *supra* note (emphasizing the rationality of participation in informational cascades).

issues of race – a change in position, after exposure to the views of others, also seems entirely rational.

The problem with group polarization is not that people subject to it suffer from some cognitive or motivational defect. The problem is instead that people may be shifted, as a result of entirely rational processes, in the direction of factual, legal, or moral mistakes.

E. The Relativity of Polarization and Polarization Framing

Notwithstanding the rationality of those subject to polarization, it should be emphasized that in laboratory studies, polarization occurs in terms of a specified issue and a specified scale. The issue for exploration is therefore framed in a certain way, and here there is a potential for manipulation. The same group of individuals, for example, might be inclined to be supportive of greater employment opportunities for women and also inclined to be skeptical about “feminism”; and polarization could drive otherwise identical groups toward more extreme positions on both questions, so much so that, in theory, one group could become very strongly committed to women’s employment opportunities that it embraced feminism, whereas another group could become so skeptical of feminism that it raised questions about greater employment opportunities for women. Here there is a lesson about the pervasive importance of “framing” in generating positions about disputed questions.⁸⁹ But there is a twist on the conventional view: In the presence of polarization, questions can be framed in such a way as to shift groups, and individuals who constitute them, in distinctive and even inconsistent directions.

Now if people attempt to square their various judgments with one another, in an attempt to reach reflective equilibrium,⁹⁰ inconsistent shifts are less likely, and people ought to be less vulnerable to framing effects. In the real world, however, it is likely that polarization occurs around issues as socially framed; cultural movements of various sorts – toward greater ethnic identification, in favor of stronger national defense, on behalf of taxpayer support for the arts – are a likely consequence. Undoubtedly political entrepreneurs, with self-interested or altruistic agendas, are in some sense aware of this fact, and attempt to produce shifts along the scale that has been made salient.

V. Implications and Illustrations

A. Outside the Laboratory

Group polarization should have a large effect on any deliberating group or institution; its effects are hardly limited to the laboratory. Religious organizations tend, for example, to strengthen group members’ religious convictions, simply by virtue of the fact that like-minded people are talking to one another.⁹¹ Indeed religious groups amplify the religious impulse, especially if group members are insulated from other groups; the result can be to lead people in

⁸⁹ See Donald A. Redermeier et al., *Understanding Patients’ Decisions*, 270 JAMA 72, 73 (1993).

⁹⁰ See John Rawls. *A Theory of Justice* (1971).

⁹¹ See David G. Myers, *Polarizing Effects of Social Interaction*, in *Group Decision Making* 125, 137-38 (Hermann Barndstatter et al. eds 1982).

quite bizarre directions.⁹² Survey evidence shows that dramatic social events, like the assassination of Martin Luther King and civil rights disturbances, tend to polarize attitudes, with both positive and negative attitudes increasing within demographic groups.⁹³ A similar process can harden attitudes toward outsiders and social change; thus proposals “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.”⁹⁴

B. Outgroups

Group polarization has particular implications for insulated “outgroups.” Recall that polarization increases when the deliberating group is able to define itself by contrast to some other contrasting group⁹⁵; outgroups are in this position – of self-contrast to others – by definition. Excluded by choice or coercion from discussion with others, such groups may become polarized in quite extreme directions, often for no better reason than group polarization. Extremism on the part of such groups (not excluding murders and suicides) is a possible result,⁹⁶ especially if we consider the fact that extreme groups show comparatively greater polarization.⁹⁷ There is also likely to be some rhetorical asymmetry within such groups, so that arguments in a certain directions have the automatic upper hand. Consider, for example, a group of people who tend to believe that academic freedom is threatened by the tendency to “political correctness” in university life; in a debate about how much attention should be paid to (say) gender studies in the curriculum, skeptics are likely to have the upper hand.

The tendency toward polarization among outgroups raises some doubts about the idea that certain group discussion produce “consciousness raising.” It is possible, at least, that the consequence of discussion is not to raise consciousness (an ambiguous term 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.⁹⁸ 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.⁹⁹

C. Feuds, Ethnic and International Strife, and War

Some of the relevant processes are at work in feuds of all kinds; one of the characteristic features of feuds is that the feuding groups tend to talk only to one another, fueling and

⁹² See Leon Festinger et al., *When Prophecy Fails* (1956).

⁹³ See R.T Riley and T.F. Pettigrew, *Dramatic Events and Attitude Change*, 34 *J Personality and Social Psych* 1004 (1976).

⁹⁴ Myers, *supra*, at 135.

⁹⁵ See Turner et al., *supra* note, at 151.

⁹⁶ Cf. Aronson, *supra* note (discussing mass suicide at Jones Beach).

⁹⁷ See *supra*; Turner et al., *supra*, at 158, 167-70.

⁹⁸ See Chip Heath and Rich Gonzalez, *Interaction With Others Increases Decision Confidence But Not Decision Quality: Evidence Against Information Collection Views of Interactive Decision Making*, 61 *Organizational Behavior and Human Decision Processes* 305 (1995).

⁹⁹ See *id.*

amplifying their outrage, and solidifying their impression of the relevant events.¹⁰⁰ Informational and reputational forces are very much at work here, producing cascade effects, and group polarization sometimes leads members to increasingly extreme positions. It is not too much of a leap to suggest that these effects are also present within ethnic groups and even nations, notwithstanding the obvious fact that here there is a high degree of heterogeneity, and deliberation cannot occur among all members at the same time.

Timur Kuran, for example, has explored the phenomenon of “ethnification.” Kuran’s basic claim is that in many nations, including Turkey and the former Yugoslavia, ethnic strife is not a reawakening of long-suppressed resentments, but instead a product 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. Hence “the fears and antagonisms that accompany high levels of ethnic activity may be a result of ethnification rather than its root cause.”¹⁰¹ 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, before discussions began. In the extreme case, the result might be war.¹⁰² And when a war begins, group polarization, if it operates at the national level, can help ensure continued hostility and antagonism.

D. The Internet and Mass Deliberation

Many people have expressed concern about processes of social influence on the Internet.¹⁰³ 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.¹⁰⁴ But an understanding of group polarization heightens these concerns and raises new ones. 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.”¹⁰⁵ If certain people are deliberating with many like-minded others, views will not be reinforced but instead 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. It is likely that processes of this general sort have threatened both peace and stability in some nations¹⁰⁶; while dire consequences are unlikely in the United States, both fragmentation and violence are predictable results. As we have seen, group polarization is intensified if people are speaking anonymously and if attention is drawn,

¹⁰⁰ See Jon Elster, *The Cement of Society* (1994).

¹⁰¹ See Timur Kuran, *Ethnic Norms and Their Transformation Through Reputational Cascades*, 27 *J Legal Stud* 623, 648 (1998).

¹⁰² See *id.* at 650-51.

¹⁰³ See Lawrence Lessig, *Code and Other Laws of Cyberspace* 186 (1999); Andrew Shapiro, *The Control Revolution* (1999).

¹⁰⁴ See *id.*

¹⁰⁵ See Patricia Wallace, *supra* note, at 73-84.

¹⁰⁶ See Timur Kuran, *supra* note, at 635-650.

through one or another means, to group membership. Many Internet discussion groups have precisely this feature. It is therefore plausible to speculate that the Internet may be serving, for many, as a breeding group for extremism.

Consider in this regard a study not of extremism, but of serious errors within working groups, both face-to-face and online.¹⁰⁷ The purpose of the study was to see how groups might collaborate to make personnel decisions. Resumes for three candidates, applying for a marketing manager position, were placed before the groups; the attributes of the candidates were rigged by the experimenters so that one applicant was best matched for the job described. Packets of information were given to subjects, each containing only a subset of information from the resumes, so that each group member had only part of the relevant information. The groups consisted of three people, some operating face-to-face, some operating on-line. Two results were especially striking: Group polarization was common; and almost none (!) of the deliberating groups made what was conspicuously the right choice, because they failed to share information in a way that would permit the group to make an objective decision. In on-line groups, the level of bias was especially high, in the sense that members tended to share positive information about the winning candidate and negative information about the losers, while also suppressing negative information about the winner and positive information about the losers. These contributions served to “reinforce the march toward group consensus rather than add complications and fuel debate,”¹⁰⁸ This tendency was twice as large within the online groups.

It is much too early to offer a confident account of the consequences of group deliberation via computer and on the Internet. But what has been said thus far should be sufficient to show that group polarization may be especially pronounced under conditions of anonymity, in a way that magnifies mistakes and biases. Though the study just described did not involve political or moral issues, the results are plausibly taken to suggest that one-sidedness, and consequently extremeness, can be heightened when communication occurs via computer.

E. Legal and Political Institutions

With respect to legal and political institutions, there is generally little direct evidence; but it is possible to venture several points.

1. Juries. Group polarization is well-documented on juries; this is the only legal institution for which direct evidence exists. In experimental settings, polarization has been found in numerous settings with respect to guilt and innocence, and indeed this appears to be an uncontradicted finding.¹⁰⁹ Outside of the experimental setting, we know that the predeliberation verdict predicts the final outcome 90% of the time, in cases where juries do not hang; this provides “powerful presumptive evidence that group polarization occurs in real juries.”¹¹⁰

As noted, a more recent study of 300 deliberating juries found massive group polarization with respect to bounded punishment scales; groups whose median pre-deliberation vote was 3 or

¹⁰⁷ See Hightower and Sayeed, *supra* note.

¹⁰⁸ Wallace, *supra* note, at 82.

¹⁰⁹ See Brown, *supra* note, at 227-29 (collecting studies).

¹¹⁰ *Id.* at 239.

less tended to generate verdicts below that of the median voter, whereas groups whose median pre-deliberation vote was above 3 tended to generate verdicts above that of the median voter. Indeed, many such juries ended up with verdicts as low or lower (for the low verdicts) as that of the lowest predeliberation voter, and as high or higher (for the high verdicts) as that of the highest predeliberation voter. I have also noted that with respect to dollars, this study did not find group polarization, at least in any simple form; positive dollar amounts generally increased, because of the rhetorical asymmetry referred to above. On the other hand: As compared to the median of predeliberation judgments, dollar amounts increased far more at the high end, and this effect is broadly consistent with the idea of group polarization.

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. These commissions have attracted attention mostly because their members are immune from plenary removal power by the President.¹¹¹ An equally striking but generally overlooked provision of the relevant statutes requires bipartisan membership: The independent commissions must be divided between Republicans and Democrats. 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’ part. An independent agency – the FCC, the NLRB, the CPSC -- 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 should 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 no direct confirmation of this general proposition. But some support comes from an intriguing study of judicial behavior on the D.C. Circuit.¹¹² Under Chevron v. NRDC,¹¹³ courts are supposed to uphold agency interpretations of law so long as the interpretations are “reasonable.” When do courts obey this stricture? The study strongly suggests that group polarization plays a role. The most important finding is a dramatic difference, on the United States court of appeals for the District of Columbia Circuit, between politically diverse panels (with judges appointed by Presidents of more than one party) and “unified” panels (with judges appointed by Presidents of only one party). On divided panels in which a majority of the court might be expected, on broadly speaking political grounds, to be hostile to the agency, the court

¹¹¹ See, e.g., Lawrence Lessig and Cass R. Sunstein, *The President and the Administration*, 94 Colum L Rev 1 (1994).

¹¹² See Frank Cross and Emerson Tiller, *Judicial Partisanship and Obedience to Legal Doctrine*, 107 Yale LJ 2155 (1998).

¹¹³ See US (1984)

deferred to the agency 62% of the time. 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. 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 (86% of the time). 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 as 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, perhaps above all because of social influence (and the importance of conveying a proper signal to fellow legislators and above all constituents). Imagine, for example, that a group of Republicans and a group of Democrats are thinking about how to vote on a proposed law – perhaps involving military spending, or an increase in the minimum wage, or mandatory parental leave legislation, or greater environmental protection. If Republicans are speaking mostly with Republicans, and if Democrats are speaking mostly with Democrats, we should expect a hardening of views toward the more extreme points. Undoubtedly this is part (certainly not all) of the explanation of party-line voting. And it is easy to imagine similar effects on Congress as a whole.

A result of this general kind has been documented with the original passage of the Clean Air Act.¹¹⁴ In the relevant period, there was a great deal of electoral pressure to enact some kind of clean air legislation. Both President Nixon and Senator Muskie attempted to signal to voters that they cared a great deal about the environment. The difficulty was that both of them found themselves in a kind of “politicians’ dilemma,” in which they had to urge more and more aggressive regulation – more aggressive, in fact, than either of them sought – precisely in order to maintain the preferred relative position vis-a-vis the electorate. Congress itself polarized accordingly, toward a more extreme position than most or even all individuals would have sought beforehand.

There are significant differences between the legislative process and the contexts in which group polarization has been studied, above all because members of Congress are subject to external political sanctions. Even if members are persuaded that a certain course of action makes best sense, they may vote otherwise, simply because of what their constituents want. Hence a limited argument pool, for members of a particular party, may matter much less than a clear signal from people back home. This point may explain occasional defections on both sides; certainly it explains why some members are able to resist both party pressures and the logic of group polarization. Unambiguous electoral signals can be a powerful buffer against that logic (though the signals themselves may be a function of group polarization within the electorate).

The same point bears on the relevance of social influence. Members of the Republican Party are likely to care a great deal what fellow Republicans think of them; but they probably

¹¹⁴ See E. Donald Elliott et al., *supra* note.

care still more about what local voters think of them. To be sure, the two are not independent of one another. If a certain Republican seems like an outlier among Republicans generally – for example, if he seems less sympathetic to the religious right than his colleagues – his electoral prospects might be damaged simply by virtue of this signal. But analytically, the two are different. Here too the votes of constituents may matter more than group deliberations (taking members of the same party as the relevant group).

It is important to underline here the fact that the mechanisms of group polarization may sometimes be at work with constituents as well. We can imagine a society in which Republicans speak mostly with each other; we can imagine a society in which Democrats speak mostly with one another too. If this is the situation, polarization should occur within political camps. We might think that group polarization supplies one of the many factors behind the sharp split between Republicans and Democrats on the impeachment of President Clinton.¹¹⁵

VI. Deliberative Trouble?

A. Doubts

The phenomenon of group polarization, alongside the phenomenon of social cascades, raises severe doubts about the value of deliberation. Note here that deliberation might be justified, as a social practice, on one of two grounds. It may be that on the question at issue, there is a truth of the matter – a correct answer – and deliberation might be justified as the best way of reaching it.¹¹⁶ Group decisions are more likely to be right than decisions made by individuals. Alternatively, we might favor deliberation for the opposite reason; 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.¹¹⁷ Group polarization raises no difficulty for the second sort of account; but it poses real problems for the first. 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? Nor do the mechanisms provide much reason for confidence. If people are shifting 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 are occurring as a result of partial and frequently skewed argument pools, the results of deliberative judgments may be far worse than the results of simply taken the median of pre-deliberation judgments.

To be sure, those who emphasize the ideals associated with deliberative democracy tend to emphasize its preconditions, which include political equality and the goal of “reaching understanding.”¹¹⁸ In real-world deliberations, behavior is often strategic, and equality is often

¹¹⁵ For more detailed discussion, see Cass R. Sunstein, *Group Dynamics*, forthcoming in *Aftermath* (2000).

¹¹⁶ This is the tendency in Gutmann and Thompson, *supra* note

¹¹⁷ See the discussion of imperfect procedural justice and pure procedural justice in John Rawls, *A Theory of Justice* (1971).

¹¹⁸ See Jurgen Habermas, *A Theory of Communicative Action* 99 (1984). Thus Habermas distinguishes between strategic and communicative action and stresses “the cooperatively pursued goal of reaching understanding”; compare the treatment in Gutmann and Thompson, *supra* note, at 52-94, emphasizing the idea of reciprocity, which emphasizes the desire to justify one’s position by reference to reasons.

absent in one or another form. But group polarization is likely to occur even in the face of equality and 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.¹¹⁹ Perhaps group polarization could be reduced or even eliminated if we emphasized that good deliberation has full information as a precondition; but that requirement is extremely stringent, and if there is already full information, the point of deliberation is greatly reduced.¹²⁰ 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.

Of course we cannot say, from the mere fact of polarization, 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 others that deserve to meet with widespread approval. Extremism need not be a word of opprobrium, and in any case a group of moderates is likely, as noted, to become entrenched in its moderation by virtue of the mechanisms discussed here. In addition, group polarization can be explained partly by reference to the fact that people who are confident are likely to be persuasive; and it seems sensible to say that as a statistical matter, people who are confident are more likely to be right. But when group discussion tends to lead people to more strongly held versions of the same view with which they began, and if social influences and limited argument pools are responsible, there is little reason for great confidence in the effects of deliberation.

B. A Lesson

As a thought experiment, 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 also limited. Social influences would undoubtedly remain; hence people might shift because of a desire to maintain their reputation and self-conception, by standing 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, in relation to the group, from what they thought it was, any shift would be in response to an accurate understanding of all relevant citizens, and not a product of a skewed group sample.

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. Perhaps weak arguments would be made and repeated and repeated again, while good arguments would be offered infrequently. But at least a deliberating body of

¹¹⁹ See Robert Goodin, *Laundering Preferences*, in *Foundations of Social Choice Theory* 75, 77-90 (Jon Elster and Aanund Hylland eds. 1986).

¹²⁰ Not eliminated. There remains the question of what to do, given a certain understanding of the facts.

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.

A possible conclusion would return to the need for full information, not only about facts but also about possible options and values, and suggest that in any deliberating body, it is important to ensure a wide mix of views, so as to ensure that a distorted argument pool does not produce unearned or unjustified shift. For a leader of any institution, it makes sense, in any ordinary circumstance, to try to ensure a broad array of views, simply in order to ensure against the predictable entrenchment of private judgments. The idea of a “public sphere” can be understood as an effort to ensure a domain in which multiple views can be heard, and can be heard by people having multiple perspectives.¹²¹ Thus there is reason for caution about any institutional practice that insulates people from competing arguments.¹²² Indeed, an understanding of group polarization suggests that it would be desirable to take steps to reduce the likelihood that panels on federal courts of appeals do not consist solely of appointees of presidents of any single political party.

Of course any argument pool will be limited; no one has time to listen to every point of view. But perhaps the largest lesson involves the need for caution about the effects of deliberation within groups all or most of whose members already have an extreme tendency. Heterogeneous groups are a far better source of good judgments. The principal qualification here is that heterogeneity is by itself neither here nor there; the 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 the view 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 they are so invidious and implausible. This point might seem to create a 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 sufficiently good sense as to generate judgments about what points of view must be included and excluded. But if we already know that, why should we not proceed directly to the merits? If we already know that, before deliberation occurs, does deliberation have any point at all?

The answer is that we often do know enough to know 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 deliberation. What is necessary is not to allow every view to be heard, but to ensue that no single view is so widely heard, and reinforced, that people are unable to engage in critical evaluation of the reasonable competitors. In this way an understanding of group polarization provides no simple view of deliberation as a social process, but does provide an important lesson to those interested in the construction of public spaces or a well-functioning public sphere.

¹²¹ See Jurgen Habermas, *The Structural Transformation of the Public Sphere* 231-50 (1991).

¹²² See Lessig, *supra* note, at 186.

Of course the provision of diverse views does not guarantee good deliberation. Among other things, most people are subject to “confirmatory bias,” in accordance with which exposure to competing position will not dislodge and may even strengthen the antecedently held position.¹²³ On questions of morality and fairness, and undoubtedly less 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.¹²⁴ But this is not a universal phenomenon, and at least an understanding of competing views is likely to weaken the forms of fragmentation and misunderstanding that come from deliberation among the like-minded.

C. The Deliberative Opinion Poll: A Contrast

In an interesting combination of theoretical and empirical work, James Fishkin has pioneered the idea of a “deliberative opinion poll,” in which small groups, consisting of highly diverse individuals, are asked to come together and to deliberate about various issues. Fishkin finds some noteworthy shifts in individual views; but he does not find a systematic tendency toward polarization. In England, for example, deliberation led to reduced interest in using imprisonment as a tool for combating crime.¹²⁵ 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%.¹²⁶ Similar shifts were shown in the direction of greater enthusiasm for procedural rights of defendants and increased willingness to explore alternatives to prison. These are not the changes that would be predicted by group polarization. The probable reason is that in Fishkin’s studies, participants were presented with a set of written materials that attempted to be balanced but that would likely move people in different directions from those that would be expected by simple group discussion. 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.¹²⁷

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%).¹²⁸ Indeed, on many particular issues, the effect of deliberation was to create an increase in the intensity with which people held their preexisting convictions.¹²⁹ These findings are consistent with the prediction of group polarization. But this was not a uniform pattern, and on some questions deliberation increased the percentage of people holding a minority position (with, for example, a jump from 36% to 57% of people favoring

¹²³ See Jonathan Baron, *Thinking and Deciding* (1994).

¹²⁴ See *id.*

¹²⁵ Fishkin, *supra* note, at 206-07.

¹²⁶ *Id.*

¹²⁷ See *id.*

¹²⁸ Fishkin and Luskin, *supra* note, at 23.

¹²⁹ 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); a jump, on a scale of 1 to 3, from 2.71 to 2.85 in intensity of support for greater spending on education; showing a jump, on a scale of 1 to 3, from 1.95 to 2.16, in commitment to aiding American business interests abroad).

policies making divorce “harder to get”).¹³⁰ Taken as a whole, a great deal of Fishkin’s data seem to support the group polarization hypothesis; what does not is probably a product of some combination of statistical noise, effects of external presentation, and deviations produced by members of the particular groups involved.

Conclusion

In this essay I have attempted to discuss the phenomenon of group polarization and to trace some of its implications for law and political theory. The basic point is that group deliberation can create polarization of both groups and individuals. The underlying mechanisms have a great deal to do with skewed and limited argument pools, and with people’s desire to maintain relative position of a certain kind (perhaps as a heuristic, perhaps for reputational reasons, perhaps because of self-conception). Group polarization can occur on juries, within legislatures, and on multimember courts and commissions. The phenomenon helps explain why many groups go, quite surprisingly, in extreme directions.

In the abstract, and without knowing about the underlying substance, it is impossible to say whether this tendency is good or bad. But the mechanisms that underlie group polarization raise serious questions about any general enthusiasm for deliberative processes. If the argument for deliberation is that it is likely to yield correct answers to social questions, group polarization suggests the need for attention to the background conditions in which this is likely to be the case. 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. I have suggested that the best response to this problem is to attempt to ensure against social balkanization and fragmentation, through mechanisms providing a “public sphere” that is used, at once, by people with competing perspectives on facts and values. If a general public sphere is unavailable or not feasible, it becomes all the more important to ensure that in the course of deliberation, people are exposed to a range of reasonable competing views.

Of course it might seem hard to know what counts as a reasonable competing view without knowing what is actually right, and if we already know that, there might seem to be little point to deliberation. But short of knowing what is right, it is possible to know something about the range of reasonable candidates, and about who might learn from whom. Perhaps the largest lesson provided by group polarization involves the need to structure processes of deliberation so as to ensure that people are exposed, not to softer or louder echoes of their own voices, but to a range of reasonable alternatives. By itself, that lesson is very far from new; but an understanding of the potential effects of group polarization argues in favor of fresh thinking, and possible reforms, in many contemporary institutions.

¹³⁰ 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 a decrease, on a scale of 1 to 3, from 2.38 to 2.27 in commitment to spending on social security).

Readers with comments should address them to:

Cass R. Sunstein
Karl N. Llewellyn Distinguished Service Professor
University of Chicago Law School
1111 East 60th Street
Chicago, IL 60637

Chicago Working Papers in Law and Economics
(Second Series)

1. William M. Landes, Copyright Protection of Letters, Diaries and Other Unpublished Works: An Economic Approach (July 1991).
2. Richard A. Epstein, The Path to *The T. J. Hooper*: The Theory and History of Custom in the Law of Tort (August 1991).
3. Cass R. Sunstein, On Property and Constitutionalism (September 1991).
4. Richard A. Posner, Blackmail, Privacy, and Freedom of Contract (February 1992).
5. Randal C. Picker, Security Interests, Misbehavior, and Common Pools (February 1992).
6. Tomas J. Philipson & Richard A. Posner, Optimal Regulation of AIDS (April 1992).
7. Douglas G. Baird, Revisiting Auctions in Chapter 11 (April 1992).
8. William M. Landes, Sequential versus Unitary Trials: An Economic Analysis (July 1992).
9. William M. Landes & Richard A. Posner, The Influence of Economics on Law: A Quantitative Study (August 1992).
10. Alan O. Sykes, The Welfare Economics of Immigration Law: A Theoretical Survey With An Analysis of U.S. Policy (September 1992).
11. Douglas G. Baird, 1992 Katz Lecture: Reconstructing Contracts (November 1992).
12. Gary S. Becker, The Economic Way of Looking at Life (January 1993).
13. J. Mark Ramseyer, Credibly Committing to Efficiency Wages: Cotton Spinning Cartels in Imperial Japan (March 1993).
14. Cass R. Sunstein, Endogenous Preferences, Environmental Law (April 1993).
15. Richard A. Posner, What Do Judges and Justices Maximize? (The Same Thing Everyone Else Does) (April 1993).
16. Lucian Arye Bebchuk and Randal C. Picker, Bankruptcy Rules, Managerial Entrenchment, and Firm-Specific Human Capital (August 1993).

17. J. Mark Ramseyer, Explicit Reasons for Implicit Contracts: The Legal Logic to the Japanese Main Bank System (August 1993).
18. William M. Landes and Richard A. Posner, The Economics of Anticipatory Adjudication (September 1993).
19. Kenneth W. Dam, The Economic Underpinnings of Patent Law (September 1993).
20. Alan O. Sykes, An Introduction to Regression Analysis (October 1993).
21. Richard A. Epstein, The Ubiquity of the Benefit Principle (March 1994).
22. Randal C. Picker, An Introduction to Game Theory and the Law (June 1994).
23. William M. Landes, Counterclaims: An Economic Analysis (June 1994).
24. J. Mark Ramseyer, The Market for Children: Evidence from Early Modern Japan (August 1994).
25. Robert H. Gertner and Geoffrey P. Miller, Settlement Escrows (August 1994).
26. Kenneth W. Dam, Some Economic Considerations in the Intellectual Property Protection of Software (August 1994).
27. Cass R. Sunstein, Rules and Rulelessness, (October 1994).
28. David Friedman, More Justice for Less Money: A Step Beyond *Cimino* (December 1994).
29. Daniel Shaviro, Budget Deficits and the Intergenerational Distribution of Lifetime Consumption (January 1995).
30. Douglas G. Baird, The Law and Economics of Contract Damages (February 1995).
31. Daniel Kessler, Thomas Meites, and Geoffrey P. Miller, Explaining Deviations from the Fifty Percent Rule: A Multimodal Approach to the Selection of Cases for Litigation (March 1995).
32. Geoffrey P. Miller, Das Kapital: Solvency Regulation of the American Business Enterprise (April 1995).
33. Richard Craswell, Freedom of Contract (August 1995).
34. J. Mark Ramseyer, Public Choice (November 1995).
35. Kenneth W. Dam, Intellectual Property in an Age of Software and Biotechnology (November 1995).

36. Cass R. Sunstein, Social Norms and Social Roles (January 1996).
37. J. Mark Ramseyer and Eric B. Rasmusen, Judicial Independence in Civil Law Regimes: Econometrics from Japan (January 1996).
38. Richard A. Epstein, Transaction Costs and Property Rights: Or Do Good Fences Make Good Neighbors? (March 1996).
39. Cass R. Sunstein, The Cost-Benefit State (May 1996).
40. William M. Landes and Richard A. Posner, The Economics of Legal Disputes Over the Ownership of Works of Art and Other Collectibles (July 1996).
41. John R. Lott, Jr. and David B. Mustard, Crime, Deterrence, and Right-to-Carry Concealed Handguns (August 1996).
42. Cass R. Sunstein, Health-Health Tradeoffs (September 1996).
43. Douglas G. Baird, The Hidden Virtues of Chapter 11: An Overview of the Law and Economics of Financially Distressed Firms (March 1997).
44. Richard A. Posner, Community, Wealth, and Equality (March 1997).
45. William M. Landes, The Art of Law and Economics: An Autobiographical Essay (March 1997).
46. Cass R. Sunstein, Behavioral Analysis of Law (April 1997).
47. John R. Lott, Jr. and Kermit Daniel, Term Limits and Electoral Competitiveness: Evidence from California's State Legislative Races (May 1997).
48. Randal C. Picker, Simple Games in a Complex World: A Generative Approach to the Adoption of Norms (June 1997).
49. Richard A. Epstein, Contracts Small and Contracts Large: Contract Law through the Lens of Laissez-Faire (August 1997).
50. Cass R. Sunstein, Daniel Kahneman, and David Schkade, Assessing Punitive Damages (with Notes on Cognition and Valuation in Law) (December 1997).
51. William M. Landes, Lawrence Lessig, and Michael E. Solimine, Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges (January 1998).
52. John R. Lott, Jr., A Simple Explanation for Why Campaign Expenditures are Increasing: The Government is Getting Bigger (February 1998).

53. Richard A. Posner, Values and Consequences: An Introduction to Economic Analysis of Law (March 1998).
54. Denise DiPasquale and Edward L. Glaeser, Incentives and Social Capital: Are Homeowners Better Citizens? (April 1998).
55. Christine Jolls, Cass R. Sunstein, and Richard Thaler, A Behavioral Approach to Law and Economics (May 1998).
56. John R. Lott, Jr., Does a Helping Hand Put Others At Risk?: Affirmative Action, Police Departments, and Crime (May 1998).
57. Cass R. Sunstein and Edna Ullmann-Margalit, Second-Order Decisions (June 1998).
58. Jonathan M. Karpoff and John R. Lott, Jr., Punitive Damages: Their Determinants, Effects on Firm Value, and the Impact of Supreme Court and Congressional Attempts to Limit Awards (July 1998).
59. Kenneth W. Dam, Self-Help in the Digital Jungle (August 1998).
60. John R. Lott, Jr., How Dramatically Did Women's Suffrage Change the Size and Scope of Government? (September 1998)
61. Kevin A. Kordana and Eric A. Posner, A Positive Theory of Chapter 11 (October 1998)
62. David A. Weisbach, Line Drawing, Doctrine, and Efficiency in the Tax Law (November 1998)
63. Jack L. Goldsmith and Eric A. Posner, A Theory of Customary International Law (November 1998)
64. John R. Lott, Jr., Public Schooling, Indoctrination, and Totalitarianism (December 1998)
65. Cass R. Sunstein, Private Broadcasters and the Public Interest: Notes Toward A "Third Way" (January 1999)
66. Richard A. Posner, An Economic Approach to the Law of Evidence (February 1999)
67. Yannis Bakos, Erik Brynjolfsson, Douglas Lichtman, Shared Information Goods (February 1999)
68. Kenneth W. Dam, Intellectual Property and the Academic Enterprise (February 1999)
69. Gertrud M. Fremling and Richard A. Posner, Status Signaling and the Law, with Particular Application to Sexual Harassment (March 1999)

70. Cass R. Sunstein, Must Formalism Be Defended Empirically? (March 1999)
71. Jonathan M. Karpoff, John R. Lott, Jr., and Graeme Rankine, Environmental Violations, Legal Penalties, and Reputation Costs (March 1999)
72. Matthew D. Adler and Eric A. Posner, Rethinking Cost-Benefit Analysis (April 1999)
73. John R. Lott, Jr. and William M. Landes, Multiple Victim Public Shooting, Bombings, and Right-to-Carry Concealed Handgun Laws: Contrasting Private and Public Law Enforcement (April 1999)
74. Lisa Bernstein, The Questionable Empirical Basis of Article 2's Incorporation Strategy: A Preliminary Study (May 1999)
75. Richard A. Epstein, Deconstructing Privacy: and Putting It Back Together Again (May 1999)
76. William M. Landes, Winning the Art Lottery: The Economic Returns to the Ganz Collection (May 1999)
77. Cass R. Sunstein, David Schkade, and Daniel Kahneman, Do People Want Optimal Deterrence? (June 1999)
78. Tomas J. Philipson and Richard A. Posner, The Long-Run Growth in Obesity as a Function of Technological Change (June 1999)
79. David A. Weisbach, Ironing Out the Flat Tax (August 1999)
80. Eric A. Posner, A Theory of Contract Law under Conditions of Radical Judicial Error (August 1999)
81. David Schkade, Cass R. Sunstein, and Daniel Kahneman, Are Juries Less Erratic than Individuals? Deliberation, Polarization, and Punitive Damages (September 1999)
82. Cass R. Sunstein, Nondelegation Canons (September 1999)
83. Richard A. Posner, The Theory and Practice of Citations Analysis, with Special Reference to Law and Economics (September 1999)
84. Randal C. Picker, Regulating Network Industries: A Look at *Intel* (October 1999)
85. Cass R. Sunstein, Cognition and Cost-Benefit Analysis (October 1999)

86. Douglas G. Baird and Edward R. Morrison, Optimal Timing and Legal Decisionmaking: The Case of the Liquidation Decision in Bankruptcy (October 1999)
87. Gertrud M. Fremling and Richard A. Posner, Market Signaling of Personal Characteristics (November 1999)
88. Matthew D. Adler and Eric A. Posner, Implementing Cost-Benefit Analysis When Preferences Are Distorted (November 1999)
89. Richard A. Posner, Orwell versus Huxley: Economics, Technology, Privacy, and Satire (November 1999)
90. David A. Weisbach, Should the Tax Law Require Current Accrual of Interest on Derivative Financial Instruments? (December 1999)
91. Cass R. Sunstein, The Law of Group Polarization (December 1999)