Cass R. Sunstein

on a danger of deliberative democracy

Imagine the following situations:

- Affirmative action is under attack in the state of Texas. A number of professors and students at a branch of the University of Texas are inclined to be supportive of affirmative action; they meet to exchange views and to plan further action, if necessary. What are these professors and students likely to think, and do, after they talk?

Cass R. Sunstein is Karl N. Llewellyn Distinguished Service Professor of Jurisprudence at the University of Chicago. A Fellow of the American Academy since 1992, he is the author of numerous articles and books, including "Free Markets and Social Justice" (1997), "Designing Democracy: What Constitutions Do" (2001), "Republic.com" (2001), and "Risk and Reason" (2002). Mr. Sunstein has testified before congressional committees on many subjects, and he has been involved in constitution-making and law reform activities in a number of nations, including Ukraine, Poland, China, South Africa, and Russia. He is currently working on various projects involving the relationship between law and human behavior.

- After a highly publicized shooting at a local high school, a group of people in the community, most of them tentatively in favor of greater gun control, comes together to discuss the possibility of imposing new gun control measures. What, if anything, will happen to individual views as a result of this discussion?
- A jury is deciding on an appropriate punitive damage award in a case of misconduct by a large company; the behavior resulted in a serious injury to a small child. Before deliberating as a group, jurors have individually considered the appropriate award, 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?

The likely behavior of individuals in these situations reveals a striking but much neglected phenomenon: that of *group polarization*. This phenomenon raises serious questions about the potential dangers of deliberation, even in some democratic settings.

In brief, the phenomenon of group polarization means that the members of a deliberating group predictably move toward a more extreme point in the direction of their pre-deliberation views.

Thus, the Texas group that meets to debate affirmative action is likely to become more firmly committed to that practice.

The community group concerned about the shooting at a local high school is likely to conclude its meeting enthusiastically in favor of gun control.

And, as a new study by David Schkade, Daniel Kahneman, and myself has shown, the jury will probably award punitive damages in excess of the median, perhaps higher than the mean as

well, and very possibly as high as or higher than the highest award selected in advance of deliberation by any individual juror.

Several factors increase the likelihood and extent of group polarization. For example, groups consisting of individuals with extremist tendencies are more likely to shift, and likely to shift more – a point that bears on the wellsprings of hatred, violence, and terrorism. The same is true for groups with some kind of salient shared identity – like Republicans, Democrats, and lawyers, but unlike jurors and experimental subjects.

It follows that when like-minded people meet regularly, without sustained exposure to competing views, extreme movements are all the more probable. Here, for example, are some empirical examples of group polarization, based on research in over a dozen nations:

- After discussion, a group of moderately profeminist American women becomes more strongly profeminist.
- After discussion, a group of French citizens becomes more critical of the United States and its intentions with respect to economic aid.
- After discussion, a group of whites predisposed to show racial prejudice offers more negative responses to the question whether white racism is responsible for conditions faced by African Americans in American cities.
- After discussion, a group of whites predisposed not to show racial prejudice offers more positive responses to the same question.

We may confidently predict, then, that those moderately critical of an ongoing war effort will, after discussion, sharply oppose the war; that those who believe that global warming is a serious problem are likely, after discussion, to hold that belief with considerable confidence; that people tending to believe in the inferiority of a certain racial group will become more entrenched in this belief as a result of discussion; that those tending to condemn the United States will, as a result of discussion, end up condemning the United States with even more intensity.

Why does group polarization occur? There are three main explanations. The first is based on *persuasive arguments*. The simple idea here is that people respond to the arguments made by other people – and that the 'argument pool,' in a group with some initial disposition in one direction, will inevitably be skewed toward that disposition. Thus a group whose members tend to think that Israel is the real aggressor in the Mideast conflict will tend to hear many arguments to that effect, and relatively few opposing views. A group whose members tend to oppose affirmative action will hear a large number of arguments in favor of abolishing affirmative action and comparatively fewer arguments for retaining it. If people are listening, they will have a stronger conviction, in the same direction from which they began, as a result of deliberation.

The second mechanism has to do with *social influence*. The central idea here is that people have a certain conception of themselves and a corresponding sense of how they would like to be perceived by others. If you think of yourself as the sort of person who favors gun control more than most people do, you might shift your position once you find yourself in a group that is very strongly in favor of gun control. If you stay where you were, you may seem less favorably disposed toward gun control than most group members, and, possibly finding your distance from the others disconcerting, you might shift more towards the group. Or if you believe that you

Note by Cass R. Sunstein

have a comparatively favorable attitude toward current policies of the Bush administration, discussion with a group whose members are at least as favorable as you might well push you in the direction of greater enthusiasm for it. Considerable evidence supports the view that social influences produce changes of this kind.

The third explanation begins by noting that people with extreme views tend to have more confidence that they are right, and that as people gain confidence they become more extreme in their beliefs. If other people seem to share your view, you are likely to become more confident that you are right. Hence it is predicted that if people learn that others agree with them, they are likely to move in a more extreme direction. In a variety of experimental contexts, reported by Robert Baron et al. in a 1996 article on "Social Corroboration and Opinion Extremity," people's views have been shown to become more extreme simply because they have been informed of the shared views of others.

In the context of punitive damage awards by juries, an especially striking phenomenon has been uncovered, one with quite general implications. Those arguing for higher awards seem to have an automatic 'rhetorical advantage' over those arguing for lower awards. The effect is so dramatic that the dollar awards of any particular jury are likely to be systematically higher than the amount chosen by the median juror before deliberation – resulting in jury awards as high as or higher than that of the *highest* individual juror in 27 percent of cases!

It is easy to imagine other contexts in which one or another side has an automatic rhetorical advantage. Consider, as possible examples, those arguing for higher penalties for those convicted of drug offenses, or those seeking to reduce tax rates. When a rhetorical advantage is involved, group deliberation will produce significant shifts in individual judgments.

Group polarization is inevitably at work in feuds, ethnic and international strife, and war. One of the characteristic features of feuds is that members of feuding groups tend to talk only to one another, fueling and amplifying their outrage and solidifying their impression of the relevant events. It is not too much of a leap to suggest that these effects are sometimes present within ethnic and religious groups and nations, even if there is a high degree of national heterogeneity. In America, sharp divergences between whites and African Americans, on particular salient events or more generally, can be explained by reference to group polarization. The same is true for sharp divergences of viewpoints within and across nations. Group polarization occurs every day within Israel and among the Palestinian Authority; it occurs within the United States and among those inclined to support, or at least not to condemn, terrorist acts. A large part of the perennial question 'Why do they hate us?' lies not in ancient grievances or individual consciences but in the social influences emphasized here.

Of course the media play a large role, simply by virtue of the arguments they repeat. It follows that if certain people are listening to stations that promote only one point of view, or reading only one set of opinions, extreme movements are possible. As I have argued in my book *Republic.com*, the phenomenon of group polarization explains why a fragmented communications market may create problems. The psychologist Patricia Wallace explains in her *The Psychology of the Internet* that 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." If certain people are deliberating with many like-minded others, views will not merely 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.

How does all this bear on the theory of democracy?

We might approach that question by noting that the framers of the American Constitution attempted to create a deliberative democracy, that is, a system that combines accountability with a measure of reflection and reason-giving. From the standpoint of political deliberation, the central problem is that widespread error and social fragmentation are likely to result when like-minded people insulated from others move in extreme directions simply because of limited argument pools and parochial influences. Compare a system of one-party domination, which stifles dissent in part because it refuses to establish space for the emergence of divergent positions; in this way, it intensifies polarization within the party while also disabling external criticism. What Irvin Janis some years ago called 'groupthink' can be understood as drawing attention to the ways in which democratic institutions can be subject to some of the same problems.

How can this be prevented? One possibility is to maintain a system of considerable diversity and checks and balances, in which different deliberating groups, subject to their own internal pressures, might reach different conclusions and ultimately correct one another's errors. In a remarkable book by an insider about America's victory in World War II (*Administrative Reflections from World War II*, by Luther Gulick), it is urged that democratic systems have a built-in advantage during war over their nondemocratic adversaries, simply because in democratic systems possible courses of action are discussed by diverse people in advance, and errors are publicized as they occur, making them more likely to be corrected.

It follows that an obvious response to the dangers of group polarization is to ensure that members of deliberating groups, whether small or large, will not isolate themselves from competing views. This point has implications for freedom of association, bureaucratic structure, and the architecture of the Internet. Indeed, the framers of the Constitution understood the system of bicameralism as a check on the risk that passions, in the form of group polarization, would lead to ill-considered decisions from one or another house. It is important to ensure that deliberation occurs within a large and heterogeneous public sphere, and to guard against a situation in which like-minded people are walling themselves off from alternative perspectives.

But there is a difficulty with this response: a certain measure of isolation will, in some cases, be crucial to the development of ideas and approaches that would not otherwise emerge and that deserve a social hearing. Members of low-status groups are often quiet within heterogeneous bodies, and thus deliberation in such bodies tends to be dominated by high-status members. A good democracy makes space for enclaves in which otherwise silent people are willing to speak and likely to be heard.

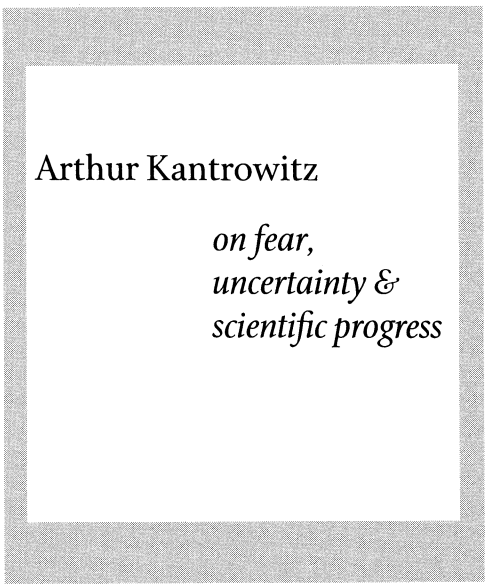
Here, then, is a dilemma: any shift – in technology, norms, or legal practice – that increases the number of deliberat-

ing enclaves will increase the diversity of society's aggregate 'argument pool' while also increasing the danger of extremism and instability, ultimately even violence.

No algorithm is available to solve the resulting conundrums.

But a simple lesson involves institutional design. To the extent that limited argument pools and social influences are likely to produce unfortunate effects, correctives can be introduced simply by exposing group members, at one point or another, to arguments to which they are not antecedently inclined. The value of deliberation, as a social phenomenon, depends very much on social context – on the nature of the process and the nature of the participants.

Here institutions are crucial. It is desirable to create spaces for deliberating groups without insulating group members from those who have opposing views, and without insulating those outside the group from the views of those within it.



In war, technological advances play a significant role. In World War II, that role was much larger than ever before, as scientists shared a great deal of the credit for the Allied victory. During and after that war, science had sufficient prestige to create a science-policy establishment in Washington powerful enough to increase federal dollars for research from

Arthur Kantrowitz, professor of engineering at Dartmouth College and retired chairman of the Avco Everett Research Laboratory, has been a Fellow of the American Academy since 1957. He is the holder of twenty-one patents, author of over two hundred scientific papers, and a member of the National Academy of Sciences and the National Academy of Engineering. He has done research in several interdisciplinary areas, including magnetically contained fusion; the invention of supersonic high-intensity molecular beams; high-temperature shock tubes, which provided the scientific basis for reentering the atmosphere from space; high-energy lasers and laser propulsion to Earth's orbit; and cardiac assist devices, which culminated in the intra-aortic balloon pump, used in millions of patients.