

Consensus and Polarization in Small Group Deliberations*

Robert C. Luskin, University of Texas at Austin

James S. Fishkin, Stanford University

Kyu S. Hahn, University of California, Los Angeles

Prepared for presentation at the annual meeting of the American Political Science Association, Chicago, IL, August 30-September 2, 2007.

ABSTRACT

One reasonable concern about discussion within social groups is that it tends to produce inauthentic consensus—not from informed, reasoned agreement but from conformism. Another, supported by jury studies and generalized by Sunstein, is that it tends to promote extremity, moving the group further out from the center of whatever dimension is at issue.

This paper examines the extent to which these tendencies occur in the ideal-ish small group deliberations in Deliberative Polling, using evidence from ten Deliberative Polls, involving 171 small groups and 59 issues. Then, after considering the extent to which the groups move toward consensus and become more extreme, we consider possible explanations for the (considerable) variation, in both respects, from group to group.

Deliberative democracy lacks for neither advocates nor critics. The latter come in three principal varieties. *Extenuationists* regard deliberation as unnecessary. People approximate their full information policy attitudes and electoral preferences without it (see, e.g., Lupia 1994). *Defeatists* regard deliberation as unattainable. People are too distracted, too intractably ignorant, too cloistered ever to achieve much in the way of inclusive, reasoned, informed and informative discussion (see, e.g., Posner 1980). *Alarmists* regard deliberation as undesirable. It may be attainable; it may affect attitudes and preferences; but it does so with certain normatively unappealing biases. Two particular allegations are that deliberation in small groups tends to homogenize opinions within groups while polarizing them across groups (see Sunstein 2000, 2002).¹

This paper addresses these two elements of the alarmist critique by examining the extent to which the ideal-ish small group deliberations (several steps toward Habermas from the real world) in Deliberative Polling homogenize and polarize opinions. We use data from ten national Deliberative Polls (DPs), varying in location, mode, and topic. One was in Australia, five in Britain, and four in the U.S. Seven were face-to-face, three online. The topics included a referendum (in Australia), a general election (in the U.K.), and presidential nomination contest (in the U.S.), as well as policy issues ranging from foreign policy to health care. These data encompass 171 small groups and 59 attitude indices.

Here we begin by considering the extent to which small groups actually exhibit these worrisome tendencies. As we shall see, the data suggest little reason for worry. There is no pronounced tendency for opinions to homogenize within groups or polarize across them. As we shall also see, however, both phenomena vary considerably from group to group. Thus we turn our attention to the question of why, estimating some simple models to explain the variation.

Deliberation and Deliberative Settings

“Deliberation’s” Latin root refers to weighing, which could in the broadest sense include solitary thought (as in Gooden). But the more common understanding, in the literature on deliberation and “deliberative democracy,” takes deliberation as involving discussion. But not just any discussion. Deliberation in this sense is discussion of a normatively desirable sort—discussion that involves a serious and at least relative unbiased weighing (the word’s Latin root) of the alternatives. Precise stipulations vary, but one reasonable set (from Fishkin and Luskin 2005) is that deliberation is discussion that is *informed and informative* (with arguments be supported by appropriate and reasonably accurate factual claims), *balanced* (with arguments met by contrary arguments, *conscientious* (with participants willing to talk and listen, with civility and respect), *substantive* (with arguments considered sincerely on their merits, not on the basis of how they are made or who is making them), and *comprehensive* (with all points of view held by significant portions of the population receiving attention).

Of course such mainstream definitions are still quite broad, leaving open many questions about the nature, institutional context, and purposes of deliberation. Who deliberates, how regularly, for how long, and under what conditions? To what extent are deliberations aimed at achieving consensus? Do the deliberations involve voting as part of the process? Is any voting public? Do the deliberations culminate in some sort of collective decision or verdict? What is the aggregation rule? (Simple majority? Supermajority? Unanimity?) With what certainty does any collective decision or verdict translate into public policy (in electoral or legislative deliberations) or judgments (in jury deliberations)? To what extent, i.e., is the outcome binding?

Different kinds of deliberation can be expected to have different effects. In every variety, deliberation should help people reach more authentic attitudes and preferences about the

subject(s) of deliberation—attitudes and preferences more in line with their own individual values and interests. In the aggregate, therefore, deliberation should promote more authentic majorities, surely important in a system founded on some version of majority rule. But overlaying and perhaps in some degree countervailing this desirable effect may be others less desirable. Sunstein (2000, 2002) has argued that deliberation generally homogenizes opinions within groups and polarizes them across groups. The people in a given discussion tend to converge on the same opinion, and one more extreme, on average, than they started with. In some cases, these tendencies presumably move people away from their more authentic attitudes and preferences, and, even if not, the widening of the gulf between opposing sides is likely to inhibit compromise, diminish civility, and worse.

These are indeed plausible tendencies, at least under certain conditions. People like to be favorably regarded by others with whom they are interacting. They seek *social approval*, and adopting the views held by those around them can help. In addition, people are influenced by the tilt of the *argument pool* to which they are exposed. Members of a group initially leaning toward one side of an issue should tend, on average, to hear more arguments on that side. If we assume that those whose views initially lie between neutrality and the initial group mean tend to move toward the latter, that those whose views initially lie at or beyond the initial group mean tend to move still further out, but that the first tendency is stronger than the second (in the sense that the movements are, on average greater), the upshot should be both increased consensus (homogenization) and increased extremity (polarization).²

But Sunstein (like other alarmist critics) takes the bulk of his evidence from jury studies, involving a very special kind of deliberation. Juries aim for consensus. Their *raison d'être* is to reach a collective decision. There is no institutional feature to ensure that the arguments on both

sides are equally considered. And the harvesting of opinions is often public. It is hardly surprising that a process designed to achieve consensus tends to homogenize opinions, nor that one lacking any means of assuring balance may produce imbalanced argument pools and centripetal tendencies.

Deliberation in other common and desirable settings may look very different. Consider the deliberation of ordinary citizens, striving to clarify their individual attitudes and electoral preferences. They cast their individual votes (and respond individually to surveys) in anonymity. The aggregate tallies reflect the knowledge and insights gained from deliberating, but each voter (or respondent) chooses independently, and there is rarely any thing approaching unanimity. Of course real world deliberation of this sort is limited and usually flawed. But it remains an important ideal for democratic theory and standard for democratic practice. How would polls read and elections come out if everyone deliberated in this fashion? What difference would it make if at least more people did?

This is the ideal that Deliberative Polling aims to approach. The participants are asked to make no collective decision, only to refine their own thinking. Their final opinions are collected only by confidential self-completion questionnaire. The discussions are led by moderators trained to enforce balance, in both participation and argument—to ensure, so far as possible, that nobody talks too much or too little and that competing arguments and counterarguments are heard. The participants are also provided with balanced information, mostly embedded in arguments for and against given policy positions. In all these respects, the small group discussions in Deliberative Polling differ dramatically from those in juries.

One other probable difference—in who deliberates—may also be worth mentioning. Both juries and DPs start with random draws, but the universe from which juries are drawn

(often, registered voters) is frequently skewed toward the better educated and more affluent. Survey houses typically start with something closer to the relevant population. On the other hand, well known biases in excuses and disqualifications in *voir dire* tend to over-represent the less well educated and less affluent on juries, while self-selection and locatability tend to have the opposite effect in Deliberative Polling samples. Which sort of sample is apt on the whole to be less biased and more heterogeneous is not entirely clear, but our money is on the Deliberative Polling samples.³

Data

The basic (face-to-face) DP design starts by interviewing a random sample, then inviting them to attend a weekend discussing the issues at a common site. Before the weekend, they are sent carefully balanced briefing materials laying out arguments for and against policy alternatives. During the weekend, they alternate between discussions in randomly assigned small groups led by trained moderators and plenary sessions in which they put questions shaped by the small group discussions to panels of experts, policy makers, or politicians. During the weekend, they alternate between discussions in randomly assigned small groups led by trained moderators and plenary sessions in which they put questions shaped by the small group discussions to panels of experts, policy makers, or politicians. At the end, the participants answer the same questions as at the beginning. Frequently a separate random sample, answering the same questions at the end of the process, provides a control group. (For more on the face-to-face design, see Fishkin 1997; Fishkin and Luskin 1999; Luskin, Fishkin, and Jowell 2002.)

The online design is broadly similar. The mode is voice, not text, so what is lost is the physical presence and company of others, as well as the nonvocal communication entailed. The discussions are spread over a period of weeks. The assignment to small groups is nonrandom,

dictated instead by possible meeting times. These are the main differences. But the participants still “meet” in small groups, still led by trained moderators, and still question panels of policy experts and policy makers. They still receive briefing materials. They are still randomly sampled offline, with “nonusers” receiving free access, including free equipment, to get them online.⁴ (For more on the online design, see Luskin, Fishkin, and Iyengar 2006 and Iyengar, Luskin, and Fishkin 2006.)

For the most part, our data come from ten Deliberative DPs, summarized in Table 1. Five were in the U.S., four in the U.K., and one in Australia. Three were online, the rest face-to-face. The topics varied widely, from crime to foreign policy to health care. Table 2 describes

[Tables 1 and 2 about here]

the 59 policy attitude indices these DPs afford. All the attitude indices have been linearly projected onto the $[0,1]$ interval, for comparability’s sake. As can be seen, there was quite a lot of net policy attitude change. This appears to be typical of Deliberative Polling (Luskin, Fishkin, and Hahn 2007). But did the attitudes polarize and homogenize, as might be feared?

Does Deliberation Homogenize and Polarize Opinions? By How Much?

For homogenization, we may look at the variance of within-group attitudes, a measure of heterogeneity. The worry here is that the variance tends to decrease. A simple way of gauging this is the percentage of groups showing a decrease, as in Luskin, Fishkin, and Jowell (2002). Across all the 171 groups on all the 59 indices in these 10 DPs, this percentage is 57.2% (significantly different from 50.0% at the .01 level). There is some tendency for opinions within groups to converge. On the other hand, that leaves 42.8% of the groups—not so many fewer than half of them—where opinions diverge. The tendency to homogenize is modest. Table 3 presents similar results for a still larger number of DPs (15) involving a still larger

(Table 3 about here)

number of small groups (263) and policy indices (107). Again the tendency to homogenize is present but modest. Roughly 56 percent of the small group-policy index combinations show a decrease in within-group variance, but that means that roughly 44 percent do not.

Perhaps a better criterion, however, is the *magnitude* of the average increase or decrease. The opinions in half the groups could have diverged, yet the deliberation still have had a homogenizing effect, to the extent that the opinions in those groups diverged less than the opinions in the other half of the groups converged. By this reckoning, too, deliberation seems to produce some but limited convergence. The mean decrease in the standard deviation, though highly significant ($p < .01$) is a mere .015.

For polarization, the question is how the group mean changes in relation to the scale midpoint. Does it tend to move further out on the same side, as Sunstein would have it, or back toward (or past) the midpoint? Across all the 171 groups on the 59 indices in our 10 DPs, the percentage moving further out averages only 52.5%, insignificantly different from 50.0%. Table 3 again presents the results for the larger set of 263 small groups on 107 policy indices in 15 DPs. Here the difference from 50.0% is still smaller (0.2%) and less significant.

Again the more relevant test may concern the average magnitude of movement, but again this test tells a similar story. Across our 171 small groups and 59 policy indices, the groups average moving, also insignificantly, .003 back *toward* the midpoint. In short, there is no evidence whatever of routine polarization. Some groups become more extreme, others more moderate, in about equal proportions and, on average, in about equal degree.

Determinants of Polarization and Homogenization

The considerable group-to-group and issue-to-issue variation in these results invites explanation. Why do given groups converge or diverge, polarize or depolarize, on given issues? Here we work with the magnitudes of these changes, pooling across groups and indices. Let us refer to these dependent variables simply as *homogenization* and *polarization*. Homogenization is defined as $h_{ig} = -(s_{ig2} - s_{ig1})$, and polarization as $p_{ig} = d_{ig}(\bar{x}_{ig2} - \bar{x}_{ig1})$, where \bar{x}_{igt} and s_{igt} stand for the sample mean and standard deviation of the i^{th} policy attitude index in the g^{th} group at time t ($= 1, 2$) and d_{ig} is a dummy variable that $= 1$ if $\bar{x}_{ig1} > .5$ (the midpoint) and $= -1$ if $\bar{x}_{ig1} < .5$ (so that change further out on the same side of the scale is always scored positively).

The obvious statistical complication in modeling the effects on these dependent variables in these data, is that the observations may well not be independent. The disturbance for different issues in the same group or (somewhat less likely) for different groups on the same issue may be correlated. Thus an appropriate statistical model is a hierarchical linear model (HLM) with two levels of cross-classified random effects (Bryk and Raudenbush, 1992, Goldestein 1987), corresponding to groups and issues.⁵ We assume the random effects of the groups and issues to be additive (as in Raudenbush and Bryk 2002, pp. 375). We also assume the disturbance to be normally distributed, all the coefficients besides the intercepts to be fixed, and the within-cell variance of the disturbance to be constant.

So what group-level attributes might affect polarization and homogenization? One likely suspect is the extremity of pre-deliberation attitudes (EXTREMITY), defined as the absolute value of the distance between the group mean and the midpoint at time 1. The more extreme a group's initial mean position, the more imbalanced its argument pool, and the more it may tend both to homogenize and to polarize (consistent with Sunstein 2000, 2002). On the other hand,

there could be something like a ceiling effect, with more extreme groups having less room to become more so.

Another strong possibility is the group's initial level of heterogeneity (DIVERSITY), defined as the group's time 1 variance. From the perspective of the argument pool, greater diversity should tend reduce both homogenization and polarization. But one could also imagine its having an opposite effect on homogenization, on a room-for-change argument. The more initially heterogeneous the group's opinions, the more room there is for it to homogenize.

Other possibilities include country, mode, and policy domain. Thus we include a dummy variable to distinguish the U.S. DPs (U.S.), a mode dummy to distinguish the online DPs (ONLINE), and a domain dummy to distinguish foreign policy issues (FOREIGN).

Table 3 shows the maximum likelihood coefficient estimates. Consider first the results for polarization, which are the weaker. Puzzlingly, diversity decreases, not increases, polarization. We are unsure of the reason. For homogenization the results are stronger and more

[Table 4 about here]

expected. Online DPs appear to have produced less homogenization than face-to-face DPs, in keeping with the less vivid, less intimate communication involved (although the coefficient estimate only approaches statistical significance, at $p=.15$). More certainly, initial extremity increased homogenization, in keeping with an argument-pool effect, while initial diversity, also increased it, in keeping with a room-for-change effect instead.

Discussion

In some institutional settings, notably including juries, deliberation may culminate in a collective decision. But in others it doesn't. In some institutional settings, again notably including juries, it may aim at achieving consensus. But again in others it doesn't. The

deliberation of ordinary citizens, striving to clarify their individual attitudes and electoral preferences, is still deliberation—of a common and important kind. Voters cast their individual votes (and respond individually to surveys) in anonymity. The aggregate tallies reflect the knowledge and insights gained from deliberating, but each voter (or respondent) chooses independently, and there is rarely any thing approaching unanimity.

Here we have used evidence from more than a dozen DPs to examine and explain the extent to which the members of randomly assigned, moderated small groups adopt more extreme views (polarize) and tend to adopt the same view (homogenize). Averaging across groups and issues, we find some modest homogenization but no polarization. The groups and issues vary considerably in these respects, however. We report some first-cut analyses of what may account for this variation. We have had limited success to this point in explaining the variation in polarization, but homogenization appears to depend on the mode of deliberation and the extremity and diversity of the group's initial views.

References

- Bryk, A. S. and Raudenbush, S. W. 1992. *Hierarchical Linear Models*. Sage Publications.
- Fishkin, James S. 1995. *The Voice of the People: Public Opinion and Democracy*. New Haven: Yale University Press.
- Fishkin, James S. and Luskin, Robert C. 1999. "Bringing Deliberation to the Democratic Dialogue." In Maxwell McCombs and Amy Reynolds, eds., *The Poll with a Human Face: The National Issues Convention Experiment in Political Communication*. Mahwah, N.J.: Lawrence E. Erlbaum.
- Fishkin, James S. and Luskin, Robert C. 2005. "Experimenting with a Democratic Ideal: Deliberative Polling and Public Opinion." *Acta Politica* 40: 284-98.
- Goldstein, H. 1987. *Multilevel Models in Educational and Social Research*. Griffin, London: Oxford University Press.
- Lupia, Arthur 1994. "Shortcuts versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections." *American Political Science Review*. 88: 63-76.
- Luskin, Robert C., Fishkin, James S., and Roger Jowell. 2002. "Considered Opinions: Deliberative Polling in Britain." *British Journal of Political Science*, 32:455-87.
- Luskin, Robert C., James S. Fishkin, and Kyu S. Hahn. 2007. "Deliberation and Net Attitude Change." Paper prepared for presentation at the ECPR General Conference, Pisa, Italy, September 6-8, 2007.
- Luskin, Robert C., James S. Fishkin, and Shanto Iyengar. 2005. Considered Opinions on U.S. Foreign Policy: Evidence from Online and Face-to-Face Deliberative Polling. Paper presented at the Annual Meeting of American Political Science Association (APSA).
- Raudenbush, S. W., and Bryk, A. S. 2002. *Hierarchical linear models: Applications and data analysis methods* (2nd edition). Newbury Park, CA : Sage.

Sunstein, Cass R. 2000. "Deliberative Trouble? Why Groups Go to Extremes." *The Yale Law Journal*, 110(1): 71-119.

Sunstein, Cass R. 2002. "The Law of Polarization." *The Journal of Political Philosophy*, 10(2): 175-195.

Table 1
Deliberative Polls Analyzed

Study	Participants	Country	Mode	Indices	Groups	Mean Group Size
Constitutional Referendum	347	Australia	F2F	5	24	14.5
General Election (1997)	276	U.K.	F2F	4	15	18.4
Europe	238	U.K.	F2F	3	16	14.9
National Health Service	230	U.K.	F2F	12	15	15.3
Monarchy	258	U.K.	F2F	4	15	17.2
New Haven	132	U.S.	F2F	3	16	8.3
Foreign Policy	340	U.S.	F2F	9	24	14.2
Nomination Contest (2004)	283	U.S.	OL	4	16	17.7
Foreign Policy (Online)	381	U.S.	OL	9	15	25.4
General Election (2004)	335	U.S.	OL	6	15	22.3
Total/Mean	2820			59	171	16.5

Table 2
Policy Attitude Indices
(Before and After Deliberation)

Study	Index	Before		After	
		Mean	SD	Mean	SD
Australia	Democracy	.606	.272	.666	.246
	National Autonomy	.635	.292	.740	.229
	Politicization	.707	.273	.469	.267
	Tradition	.622	.294	.569	.272
	Workability	.577	.241	.640	.258
British General Election (1997)	E.U.	.461	.351	.591	.290
	Minimum Wage	.789	.303	.734	.310
	Redistribution	.687	.294	.743	.235
	Tax	.808	.197	.713	.238
Europe	East European Expansion	.548	.221	.538	.266
	Favor a Referendum	.741	.235	.802	.253
	Scope of EU	.490	.270	.489	.246
National Health Service	Cutting Expensive Treatments	.634	.226	.606	.213
	Discretion	.770	.198	.708	.201
	Doctors Discretion	.774	.197	.779	.200
	Healthcare Funding Options	.326	.207	.344	.211
	Healthcare Payer	.818	.259	.888	.209
	Healthcare Options	.869	.232	.886	.197
	Patients have more say	.718	.188	.761	.151
	Priority to Poor	.616	.292	.549	.313
	Privatizing Some Treatments	.436	.251	.389	.279
	Priority to Prevention	.619	.302	.587	.274
	Quality of Life	.480	.295	.533	.295
	Waiting List	.550	.431	.649	.368
	Monarchy and Populism	.589	.171	.624	.178
	Power of Monarchy	.400	.311	.431	.314
	Reform HOL	.599	.366	.645	.298
	Support for Monarchy	.664	.317	.635	.290
New Haven	Expanding vs. Ending Airport Service	.770	.253	.708	.275
	Mandatory vs. Voluntary Revenue Sharing	.417	.143	.305	.170
Foreign Policy	Fighting Poverty and Suffering	.588	.181	.682	.165
	Fighting Terrorism	.802	.160	.822	.136
	Human Rights	.701	.216	.727	.181
	Increasing Foreign Aid	.479	.327	.731	.306
	Internationalism	.746	.286	.847	.241
	Liberalizing Trade	.613	.392	.578	.388
	Multilateralism	.733	.130	.785	.094
	Protecting the Environment	.777	.203	.770	.208
	Promoting Democracy	.632	.191	.686	.166

Presidential Nomination (2004)	Multilateralism	.696	.276	.773	.233
	Defense vs. Domestic Issues	.609	.231	.635	.209
	Free Trade	.639	.235	.676	.257
	Services vs. Taxes	.638	.222	.649	.226
Foreign Policy (Online)	Democracy	.511	.164	.535	.147
	Fighting Poverty and Suffering	.449	.149	.477	.135
	Fighting Terrorism	.805	.133	.798	.139
	Human Rights	.583	.234	.630	.192
	Increasing Foreign Aid	.314	.322	.415	.347
	Internationalism	.669	.289	.693	.282
	Liberalizing Trade	.352	.288	.398	.282
	Multilateralism	.731	.140	.746	.122
	Protecting the Environment	.682	.228	.712	.229
	Constitutional Rights	.579	.395	.523	.349
General Election (2004)	Free Trade	.688	.323	.658	.310
	Gay Marriage	.665	.419	.641	.429
	Health Insurance	.473	.390	.527	.391
	Multilateralism	.679	.368	.643	.361
	Services	.591	.353	.550	.305

Table 3
Polarization and Homogenization across 15 Deliberative Polls

Poll	No. of Small Groups	No. of Indices	Proportion of small group means moving away from midpoint	Proportion of small groups with decreasing variance
Australia	24	5	0.525	0.617
British Crime	20	5	0.460	0.530
British Europe	16	4	0.563	0.328
British General Election	15	4	0.383	0.650
British Health	15	13	0.518	0.564
British Monarchy	15	4	0.567	0.567
Bulgaria I	17	13	0.394	0.511
China	16	10	0.475	0.706
Greece	15	6	0.333	0.233
New Haven	16	3	0.479	0.542
NIC I	26	7	0.582	0.637
NIC II	24	9	0.713	0.694
Electric Utilities: CPL	16	8	0.594	0.625
Electric Utilities: SWEPCO	14	8	0.518	0.598
Electric Utilities: WTU	14	8	0.429	0.634
Mean	17.5	7.1	0.502	0.562

Table 4
Determinants of Polarization and Homogenization

	Polarization			Homogenization		
Constant	-.047	(.020)	***	-.192	(.014)	***
U.S.	.029	(.027)		.015	(.023)	
ONLINE	-.035	(.026)		-.034	(.021)	
FOREIGN	.016	(.022)		.019	(.018)	
EXTREMITY	.036	(.036)		.092	(.021)	***
DIVERSITY	.109	(.055)	**	.726	(.032)	***
Log-likelihood	-914.4			-1534.1		
<i>N</i>	1021			1021		

* $p < .10$; ** $p < .05$; *** $p < .01$. Note: Standard errors in parenthesis.

*We thank Alice Siu, Michael Weiksner, and Jennifer McGrady for research assistance.

A slightly earlier version of this paper was presented at the annual meeting of the Midwest Political Science Association, Chicago, IL, April 12-15, 2007.

¹A third, not addressed here, is that deliberation tends to move opinions in the direction of those held by the socially advantaged (see Young xxxx and Sanders xxxx).

²Sunstein also raises the possibility that the collective decision making fosters risk acceptance and thus extremity.

³To be fair, Sunstein (2000, 2002) acknowledges our earlier results, based on the first DP, in Britain in 1994, showing no great evidence of polarization (see Luskin, Fishkin, and Jowell), and cites the balance of the DP's discussions and the heterogeneity of its participants as possible reasons.

⁴All three online DPs relied on the online polling firm Knowledge Networks for sampling and recruitment.

⁵*Cross-classified* are to be distinguished from *nested* random effects. In the former, each lower level unit (group-issue combination, in our case) falls under one and only one unit at the next higher level; in the former, each lower level unit falls under a combination of two or more different kinds of higher level units (groups and issues, in our case).