Simulating the Ideal eDeliberation: The Roles of Inclusion, Equalization and Rationalization

Weiyu Zhang

Communications and New Media
National University of Singapore, Block AS6, Level 3,
11 Computing Drive, Singapore, 117416
cnmzw@nus.edu.sg

Abstract. eDeliberation refers to an emerging body of practices that purposely foster open, fair, and rational discussions over the Internet. However, the ideal concept of deliberation is confronted with the complex social conditions, such as passive citizenship and participatory inequalities. Simulation modeling was used to create situations in which (1) all people who were invited to participate in two eDeliberation projects actually attended the discussions (i.e., the openness-inclusion scenario), (2) all people equally expressed their opinions and supported their opinions with reasons (i.e., the fairness-equalization scenario), and (3) all people expressed their opinions vigorously and provided maximum arguments to support their opinions (i.e., the rationality-maximization scenario). By comparing the observed after-deliberation opinion distributions with these simulation results, we can see how public opinion generated from an ideally inclusive, equal, and argumentative procedure of deliberation will be different. Simulation findings suggest that 44% of comparisons show sizable differences. Rationaliy-maximization has the strongest impact on opinion distributions. Inclusion has relatively modest influences on opinion changes. Equalization, unexpectedly, has no influence on most opinion measures.

1. Introduction

Deliberation is a communication procedure that is open, fair, and rational [1]. Unfortunately, most of our day-to-day communication does not fit these criteria. Taking discursive participation as an example, everyday political talk between family members and friends is not open enough to include diverse opinions [2]; call-in radio discussions are open but not always reason-centered, along with a dominant role of the host [3]; opinion polls may be open (if the randomness of samples is achieved), fair (when questions make the same sense to every respondent), but not necessarily reason-centered (because respondents do not have to appeal to their rationality to give an answer) [4]. eDeliberation takes advantage of the Internet to engage ordinary citizens in open, fair and rational discussions. It is claimed that the Internet and its ability to transcend the time and space limits make some of the constraints uncritical [5]. For example, the homogeneity that is often associated with everyday political talk will not be the case when diverse people can meet online [6].

However, deliberation practices, including eDeliberation ones, are confronted with the complex social conditions in which they have to operate. These social conditions, such as structural inequalities [7] and passive citizenship [8], might render practices unable to fulfill the ideal of deliberative communication. An open procedure might not be able to lead to universal participation due to the lack of resource to support such participation. Giving participants equal opportunity to voice their opinions does not necessarily mean that everyone will take the chance, because there exist various motivation and resource concerns. Although rationality is central to deliberation, the questions that are supposed to elicit reasonable arguments do not always obtain rational responses. Personal tangents and emotional expressions also appear in deliberate discussions. The persistence of realistic social constraints raises a doubt about the deliberation practices: Are the results of deliberation legitimate when the procedure does not fit the ideal perfectly?

Empirical examination can help us to answer this question by comparing the observed results of deliberation to those which might have been generated in an idealized situation, namely, a fully inclusive, absolutely fair, and highly argumentative procedure. This paper attempts to do two things: First, by the aid of simulation modeling, the consequences of an idealized procedure can be simulated. Second, through the comparison between what is observed and what is simulated, differences can be seen and judgment regarding the legitimacy of deliberate decisions can be made.

2. Method

2.1 Data

The data come from the Electronic Dialogue 2000 project (ED2K) and the Healthcare Dialogue project (HCD)[1], two multi-wave panel projects each lasting roughly one year. The two projects are distinguished from other deliberation studies and the Internet-based studies in a number of ways. While most deliberation studies examine deliberative practices in a face-to-face setting [9], ED2K and HCD take advantage of the unique capacities of the Internet and World Wide Web for circulating information, conveying public discourse, and gathering survey data. Different from most Internet-based studies [10], which examine asynchronous message boards or less formal and happenstance “chat” experiences on the Web, both projects here created synchronous, real-time, moderated group discussions that were designed specifically to produce useful citizen deliberation. Facilitation/moderation was present and, more importantly, was standardized across both discussions and groups. In addition, neither project relied on a convenience sample of Internet users, as is common in most deliberation studies and Web-based studies. Instead, they began with a broadly representative sample of Americans and attempted to recruit from that sample a set of discussion groups that would be, in their entirety, as nearly representative as possible of U.S. citizens. In order to address the digital divide concern, all the people included in the sample were offered free equipment, free Internet, and free training, if needed.

[1] Principal Investigators on both projects are Vincent Price, Ph.D., The Steven H. Chaffee Professor of Communication and Public Opinion, and Joseph N. Cappella, Ph.D., The Gerald R. Miller Professor of Communication, both of the Annenberg School for Communication, University of Pennsylvania, USA. The findings only represent the author’s opinions.

The core of both projects consisted of groups of citizens who engaged in a series of real-time electronic discussions about issues facing either the unfolding 2000 presidential campaign or the country's healthcare reform. A set of baseline surveys assessed participants' opinions, communication behaviors, political psychology, political activities, and a variety of demographic, personality, and background variables. Subsequent group deliberations generally included pre- and post-discussion surveys. The full text of all group discussions, which lasted an hour apiece, was recorded. A series of end-of-project surveys were then conducted after the last discussion was finished. This paper utilizes two types of data: surveys and discussion transcripts. The surveys included recruitment, baseline, post-discussion, and end-of-project surveys. Content analysis was carried out on discussion transcripts to measure the amount of talk and arguments during eDeliberation

2.2 Simulation modeling

Simulation here refers to the methodology of creating an artificial representation of a real world system in order to manipulate and explore the properties of that system [11]. Simulation as a methodology has not been fully recognized in communication research. The majority of simulation studies we can see in communication research are actually either computer or statistical simulations, which are distinct from the modeling method discussed here. However, simulation actually fits the need of communication research and opens up the possibility of predicting complicated communication trends. Not all modes of communicative actions can be readily observed and analyzed in the reality. Simulation methods provide us a tool that can test even the most idealist modes of communication and their influence.

The fundamental question that simulation modeling tries to answer is – What if? For example, what if group members interact with each other in a perfectly fair situation? Challenges about the preciseness of these answers are always legitimate because simulation is highly constrained by the modeling assumptions. However, a significant strength of simulation is that everything is open to adjustment. For example, if one thinks that group members should not be equally talkative and rather randomly eloquent, we can definitely change the distribution of the amount of talk variable and then simulate the products. What might be more fruitful is to first determine which products we want to see and then go back to change possible functioning variables. For instance, if we want to see a consensus among group members, we can change either the demographic composition of groups, or the communicative procedure, or the initial opinion distributions. We can compare all these possible controls and choose those that are most promising in current situations as guidelines for intervention.

2.3 Procedure

Simulation involves a set of important assumptions. In addition to the assumptions of data missing at random, accurate model specification, and accurate coefficients, simulation assumes that changing the distributions of certain predictor variables (i.e. amount of talk and number of arguments) does not change their relationships with other variables in the model. Specifically, both the coefficients and the distributions of other variables remain the same, despite the fact that the distributions of particular variables in concern have been altered.

Following the logic discussed above, simulations in this chapter went through steps that are very similar to those used by Althaus [12]. In the first step, all opinion

and policy preference questions were recoded into dummies: “1” means supporting while “0” means not supporting. Surveyed post-discussion opinions were regressed on the demographic variables, along with one influence variable (either amount of talk or number of arguments), the pre-discussion measure corresponding to the dependent variable (missing values were imputed), and other available variables. These regression models show that the deliberation variables sometimes predict individual level post-discussion opinions (Amount of talk: 4 out of 30 ED2K measures and 3 out of 15 HCD measures; Number of reasons: 2 out of 30 ED2K measures and 3 out of 15 HCD measures). They provide support for the expectation that simulation findings might be different from observed findings. In addition, by estimating the relationships between post-discussion opinions and each of the predictors, this step provided a set of regression coefficients that can be used to simulate each person’s post-discussion opinions. These coefficients were used to model the probability that a particular individual would choose certain response alternatives to questions posed after discussions. The simulation models often have modest model fits, ranging from .02 to .14. Thirty-eight percent of the model fits were equal to or lower than .05. The mean model fit is .07.

After obtaining the coefficients for each predictor, the second step, the key step of simulation modeling, was taken. In this second stage, the what if question emerges: What if we change the distributive pattern of the deliberation variables? Which kind of consequences would we see in terms of post-discussion collective opinion distributions? Alternatively, the question could be posed this way: If we want to change the collective distributions of certain opinions, which component of the deliberation structure should we focus on? Inclusion, equalization, or maximization of influence?

This second step opens up many possible manipulations of communication procedure. This chapter examines three possibilities (see Table 1): First, the openness-inclusion scenario includes every potential participant in the deliberation regardless of their different backgrounds, assigns these potential participants the mean values of deliberation variables, and examines the difference between simulated all’s and observed attendees’ opinions. Second, the fairness-equalization scenario relies on actual attendees, but uses the means of deliberation variables rather than the observed values for each attendee who did voice his or her opinions and compares the simulated attendees’ opinions to those actually observed. Third, the influence-maximization scenario relies on actual attendees, but changes the influence values into either the highest or the mean scores and compares these two sets of simulated values to see whether maximization of influence makes a difference. In the third scenario, high-value simulations are compared to mean-value simulations in order to control for the equalization effect and isolate the maximization effect.

In each of these scenarios, step two involves changing each potential respondent’s score on amount of talk or number of arguments to either the highest possible value or the mean value by either replacing (if measured values are available) or imputing (if measured values are not available). In ED2K, for example, the highest possible value on the amount of talk scale was 834.50. Each potential respondent’s predicted opinions are calculated by plugging the coefficient values obtained from step one into the new models, substituting only the new values of the altered amount of talk or number of arguments variable. This step produces, for each individual, a new set of probabilities for each response alternative that simulate the

opinions every person might report, were she or he to talk as much as possible or talk at a mean level. This step relies on the 45 regression models (30 in ED2K and 15 in HCD) obtained in the first step and uses 135 simulation formulas (45 opinion measures * 3 scenarios) to exhibit the differences between simulated opinions and observed opinions.

Table 1. Theoretical models to compare simulated and comparison opinions

	Simulated Opinions	Comparison Opinions
Openness-Inclusion	Everybody	Attendees only
	Mean imputation of deliberation variables for non-attendees; Observed values for attendees	Observed values for attendees
Fairness-Equalization	Attendees only	Attendees only
	Mean imputation of deliberation variables for attendees	Observed values for attendees
Rationality-Maximization	Attendees only	Attendees only
	Maximum imputation of deliberation variables for attendees	Mean imputation of deliberation variables for attendees

The final step aggregates all of the individual simulated opinions together, including those of people who originally were missing of the responses and those who did not attend the discussions, by taking the mean of the individual probabilities for each of the alternative responses. These average probabilities, which represent collective post-discussion opinions controlling for individual differences in either amount of talk or number of arguments, will be then compared to the actual percentage supporting certain policies to reveal the differences. Statistical tests of significance of these differences are not applicable here, because simulated data involve alteration of the distributions of the predictor variables. The intent is to compare changes across a large set of opinion measures to identify some general tendencies.

3. Results

In general, 60 (42 in HCD and 18 in ED2K) out of 135 (45 opinion measures* 3 scenarios) simulated opinions differ from observed opinions at a rate equal to or higher than 5%.

3.1 Openness-inclusion

The fist comparison is between the observed opinion distributions and the simulated opinion distributions in an ideal scenario, in which everybody we contacted actually attended the discussions and either talked or argued at a mean level of amount. In ED2K, 13 out of 30 opinion measures examined show changes that are equal to or higher than 5%. The changes preferentially go toward more governmental

interventions, such as spending more money on health care or social security, and toward more conservative views on social issues, such as banning abortion. Some of the changes are as high as 10%. For instance, if we had all our potential participants join the discussions and be typically active, we would see that at the end of ED2K, more than half of participants (53% based on amount of talk and 55% based on number of arguments) would favor the government actions on making sure that public school students can pray as part of some official school activity, compared to a minority support (44%) actually observed.

Despite somewhat socially conservative tilt, when it comes to evaluations of presidential candidates, there is a consistent pattern showing that if we could gather full attendance, we would see significant decreases in Bush's evaluations after discussions. One of the evaluation items, viewing Bush as honest, would decrease as much as 10 percent (10% based on amount of talk and 12% based on number of arguments). In contrast, Gore's evaluation on two items would increase after discussions and one of them, making the respondent feel enthusiastic, would increase at an exceptionally high rate (17%).

In HCD, opinion measures are mainly confined to health-related policies. Here, 3 out of 15 measures show a change of preference that is equal to or higher than 5%. These changes include people's preferences on limiting drug manufacturing costs, the perceived importance of personal costs in drug policy making, and the perceived importance of tax increases in drugs policy making. The significant changes that inclusion makes suggest that descriptive under-representation of the disempowered has consequences. Descriptive under-representation can sometimes threaten the representation of opinions measured after deliberation.

3.2 Fairness-equalization

The second comparison is between the observed opinion distributions and the simulated opinion distributions in an ideal scenario, in which everybody who actually attended our discussions were equally active—either spoke an equal amount of words or provided an equal number of arguments. This scenario only produced a few changes in opinion distributions. Three out of 30 ED2K measures and 2 out of 15 HCD measures show differences that are equal to or higher than 5%. The patterns generally mirror those obtained in the first scenario. The ED2K measures show an increased positive evaluation on Gore and the HCD measures show an increased preference on limiting drug manufacturing costs and perceived importance of tax increases in drugs policy-making after discussions. Equalization, unexpectedly, has no influence on most opinion measures. It suggests that making everybody produce the same amount of words or the same number of arguments does not necessarily change opinion distributions. We might conclude that the opinion results from the two deliberation projects would not be much different were all potential participants equally argumentative.

3.3 Rationality-maximization

The third comparison is between two simulations: One is the simulation with mean values of deliberation variables among attendees and the other is the simulation with maximum values of deliberation variables among attendees. This comparison is intended to demonstrate a third scenario in which attendees either were very talkative or provided many reasons. The rationality-maximization effect is so strong that almost every variable that was examined shows a change that is equal to or

higher than 5% (26 out of 30 in ED2K, 13 out of 15 in HCD). In ED2K, different from both previous scenarios, reason-giving often leads to a decreased support in governmental interventions. For example, the support for government's financial investment in universal health care decreases almost 18% no matter which influence variable is used. Again, evaluations of presidential candidates also manifest a pattern that is quite different from those we see in the openness-inclusion and the fairness-equalization scenario. There is a consistent pattern that a highly argumentative group of attendees would have given better evaluations for Bush and lower evaluations for Gore after discussions.

In HCD, simulations show contradictory findings compared to the previous two scenarios. A rationality-maximization simulation shows at least 15% decrease in favorability toward limiting manufacturing expenses. The two concerns showing increases in the previous two scenarios, namely, personal cost and tax increases, actually show decreases (ranging from 5% to 26%) in this scenario. The other significant changes include decreases in the perceived importance of whether health care reforms would expand the size of government, are feasible, would affect the freedom to make medical decisions, would cause partisan disagreement, or affect the economy.

Reason-giving has the strongest impact on final opinion distributions. But serious questions must be resolved before we draw further conclusions: Is high rationality what we want? Furthermore, is a number-of-argument form of rationality what we want? Deliberative democracy theories answer the first question with a clear yes and with a not-so-clear answer to the second question. Habermas' communicative rationality [13] provides a different angle to look at the manifestation of rationality. Instead of defining rationality as potential persuasive influence, Habermas emphasizes mutual understanding and rationally motivated agreements. Unfortunately, the analyses in this project have to be limited to just one — and arguably not a very strong — indicator of "rationality."

3.4 Talk vs. Reasons

The last comparison is between simulations based on amount of talk vs. number of arguments. People who are most talkative do not necessarily have to be the most argumentative. Although amount of talk is often correlated with number of arguments (ED2K total Pearson correlation = .57, $p < .001$; HCD D4 Pearson correlation = .88, $p < .001$) and thus most of time the simulation findings based on the two deliberation variables are consistent in directions, we can see some interesting instances in which different deliberation variables influence outcomes in different directions.

The occasional discrepancy suggests that the effect of the amount of talk is often the same as the effect of the number of arguments. It seems that in the current deliberation practices, when people talk more, they often argue more. However, the few instances of large differences suggest that talk and argument do not always lead opinions toward the same conclusion. The explanation might be that in these instances, people do not necessarily argue more when they talk more. They might spend their eloquence on emotional expression or personal tangents, which are supposed to function differently in influencing opinion distributions. Whether this interpretation is correct is unclear, however, and cannot be resolved with the data at hand.

4. Conclusions and discussions

Simulation findings suggest that 44% of collective opinions that are predicted by simulation models differ from the observed post-discussion opinion distributions at a rate equal to or higher than 5%. In other words, if our deliberation practices were able to reach a normally ideal situation in which deliberation is fully inclusive, absolutely equal, and highly argumentative, we would see opinion results that are different from those observed. If realistic constraints prevent practices from being ideally deliberate, how much should we rely on decisions that are generated from deliberation to inform policy-making? The suggestion would be that we should treat deliberation findings as only one indicator of deliberate opinions, subject to various errors. Therefore, when we try to utilize deliberation findings to inform policy-making, we should always make clear the sources of these errors (e.g., representation of participants) and the potential size of these errors.

In addition, varying the three components leads to opinion changes in different directions. Whereas both inclusion and equalization lead to changes in the same direction, maximization of rationality often leads in an opposite direction. This contradiction implies that normative criteria of deliberation are not empirically consistent. Deliberation as a model of democracy summons forces that stretch public opinions in different directions. Thus, deliberate opinions are thus more complicated than knowledgeable opinions or informed opinions. The prediction of deliberate opinion changes is thus harder than we might expect.

All the findings above should be interpreted along with the awareness of the limitations of the simulation modeling method. The accuracy of the opinion changes predicted by the simulation models is limited by the explanatory power of the models (i.e., the model fits). Most of the simulation models in this chapter have R-square values that are low to modest in size. This is mainly because there are only a few predictor variables available for analyses. We should expect that as the number of predictors increase, we will see better model fits. A second methodological issue that is worth mentioning is that the two deliberation variables, amount of talk and number of arguments, are not always significant when used to predict individual-level post-discussion opinions. However, results are presented at the collective-level, and thus, those opinion changes that are equal to or higher than 5% do not necessarily mean that the two deliberation variables significantly predict individual opinions in those models. On the other hand, if we have significant deliberation variables at the individual level, it is certain that collective-opinion changes are significant as well. A third issue is that, in order to control for pre-discussion opinions, imputed pre-discussion opinion variables were used in the models because many cases are missing on pre-discussion measures as well. This kind of two-step modeling (the first is to impute pre-discussion opinions based on demographics and other variables, and the second is to simulate post-discussion opinions based on demographics and other variables) introduces more uncertainty into the final findings. However, since the conclusions are all about general patterns rather than specific changes, the tolerance of inaccuracy is relatively high in this set of analyses.

In summary, simulation modeling in this paper helps to provide some general predictions regarding an ideal deliberation. An ideal deliberation does probably generate collective opinions that are different from the ones observed. Openness, fairness, and reason-giving each appear to play a distinctive role in defining the ideal situation and exert idiosyncratic influences on resulting opinions. The many

significant findings in the openness-inclusion and rationality-maximization scenarios suggest that future deliberation practices should address the issues of unequal attendance and shallow rationality. However, the lack of consequences of fairness-equalization implies that unequal influence might not be as harmful as we might expect.

References

James Bohman, Williams Rehg, Deliberative democracy: Essays on reason and politics, Cambridge, MA, 1997.

Diana Mutz, Hearing the other side: Deliberative versus Participatory democracy, Cambridge, UK, 2006.

Katherin H. Jameison, Joseph N. Cappella, Echo chamber: Rush Limbaugh and the conservative media establishment, 2008.

Adam J. Berinsky, Silent voices: Public opinion and political participation in America. Princeton, NJ, 2004.

Lincoln, Dahlberg, The Internet and democratic discourse: Exploring the prospects of online deliberative forums extending the public sphere, in: Information, Communication & Society, 4/4/2001, S. 615-633.

Jennifer, Stromer-Galley, Diversity of political conversation on the Internet users' perspectives, in: Journal of Computer-Mediated Communication, 8/3/2003, retrieved on November, 29, 2006 from: http://jcmc.indiana.edu/vol8/issue3/stromergalley.html.

Sydney Verba, Katherin L. Schlozman, Henry E. Brady. Voice and equality: Civic voluntarism in American politics, Cambridge, MA, 1995.

Morris P. Fiorina, Extreme voices: A dark side of civic engagement, in: Skocpol & Fiorina (eds.), Civic engagement in American democracy, Brookings Institution Press, Washington, D. C., 1999.

James S. Fishkin, The voice of the people: Public opinion and democracy, New Haven, CT, 1995.

Eric M. Uslaner, Trust, civic engagement and the Internet, in: Political Communication, 21/2004, S. 223-242.

Thomas B. Pepinsky, From agents to outcomes: Simulation in International Relations, in: European Journal of International Relations, 11/3/2005, S. 367-394.

Scott L. Althaus, Collective preferences in democratic politics: Opinion surveys and the will of the people, Cambridge, UK, 2003.

Jurgen Habermas, The theory of communicative action, Volume 1: Reason and the rationalization of society, Boston, MA, 1984.