EERI

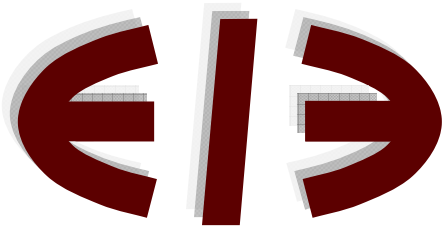
Economics and Econometrics Research Institute

Participatory Decision Making: A Field Experiment on Manipulating the Votes

Paolo Spada and James Raymond Vreeland

EERI Research Paper Series No 19/2010

ISSN: 2031-4892



EERI
Economics and Econometrics Research Institute
Avenue de Beaulieu
1160 Brussels
Belgium

Tel: +322 299 3523
Fax: +322 299 3523
www.eeri.eu

Participatory Decision Making: A Field Experiment on Manipulating the Votes

Paolo Spada[1]
and
James Raymond Vreeland[2]

Abstract: Many believe that deliberative democracy, where individuals discuss alternatives before voting on them, should result in collectively superior outcomes because voters become better informed and decisions are justified using reason. These deliberations typically involve a moderator, however, whose role has been under-examined. We conduct a field experiment to test the effects moderators may have. Participants in a class of 107 students voted on options over their writing and exam requirements. Before voting, they participated in group discussions of about five people each with one moderator. Some (randomly assigned) moderators remained neutral throughout, while others made limited interventions, supporting a specific option. We find a substantial moderator effect. Our experiment is structured like deliberations used world-wide to make community decisions and thus should have some external validity. The results indicate that if organized interest groups had influence over moderators, they might be able to hijack a deliberative decision-making process.

[1] Yale University, Department of Political Science, 124 Prospect Street, New Haven, CT 06520-8301, USA. E-mail: paolo.spada@yale.edu.
[2] Georgetown University, School of Foreign Service, 37 & O Streets, NW, Washington, DC 20057, USA. E-mail: jrv24@georgetown.edu.

1. Introduction

Many researchers argue that deliberative democracy, where individuals discuss alternatives before voting on them, should result in collectively superior outcomes because voters become better informed and their decisions are justified using reason. In most practical applications of deliberative democracy, the participants are divided into small groups of five to ten people, and a moderator is assigned to each group to facilitate a discussion or "deliberation." The small groups enable all participants to express their views. Moderators usually have the power to interrupt the discussion, asking people to intervene and present their opinions. Our question is the following: What effect might moderators have on individuals participating in group discussions?

The question is an important one because the use of participatory democracy is growing around the world. Probably the most famous example is participatory budgeting in Porto Alegre, Brazil. Every year since 1989, citizens have been called on to deliberate over various proposals and to rank the priorities that will guide the investment plan of the city. The impact of this approach on public policy – in terms of participation, redistribution, poverty reduction, and the efficiency of public spending – has been well documented (Marquetti 2003, 2007, World Bank 2008) and has led many cities in Brazil and, indeed, around the world to adopt similar practices. Non-governmental organizations (NGOs) and municipal governments in all sorts of countries – including the developed world – have also adopted the deliberative approach. Further, the process is advocated and employed by the Center for Deliberative Democracy at Stanford University, which has developed a trademarked approach called Deliberative Polling®.

The use of participatory and deliberative approaches has become particularly prevalent in developing countries where the World Bank along with the International Monetary Fund (IMF) and the United Nations (UN) have encouraged its use to make public policy decisions at the local community level. According to estimates by Mansuri and Rao (2004), World Bank lending for

"community driven development" projects grew from $325 million in 1996 to $2 billion in 2003.[3] The goals of community driven projects include the empowerment of poor people and the building of social capital (Mansuri and Rao 2004, Winters 2007). The deliberative approach has thus been employed to give a voice to all members of a community as well as to educate them.

Despite the enthusiasm for this approach to policy-making, critics warn that community-based approaches may be vulnerable to capture by local elites (Platteau and Gaspart 2003). Regarding the case of Brazil, for example, scholars have shown that the preferences of the coalition in charge of the process impacted the outcome (e.g., Baierle 2007, Wampler 2007), though ways in which elites may have hijacked the process have not been fully explored. Participatory and deliberative democracy may be manipulated in various ways. One under-explored mechanism is through the influence of moderators.

Ideally, deliberative democracy involves moderators who act as neutral referees, preventing a subgroup of participants from monopolizing the discussion. Previous research suggests, however, that moderators may actually influence the outcomes of discussions. Fulwider (2005) finds that the outcomes of deliberations with moderated groups differ from those of un-moderated groups. Similarly, Humphreys, Masters, and Sandbu (2006) find that group responses are correlated with the preferences of moderators. No previous study, however, has been designed to identify the effect of moderators on the direction of opinion change.

Can moderators influence participant preferences? We conducted a field experiment to address this question. The participants, who were unaware that they were being subjected to an experiment, made decisions in a natural environment that affected their real lives (this contrasts with lab experiments, where participants pretend they are behaving as in real life but know that an

[3] For experimental work on community driven reconstruction in the post-conflict setting of Liberia, see Fearon, Humphreys, and Weinstein (2009).

experiment is being conducted; see Gerber and Green 2008). The design was similar to real-world applications of participatory and deliberative approaches, and our subjects had a real stake in the outcome. The experiment took place in a political science class of 107 students who participated in two referenda – one on the course writing requirements and the other on the course exam requirements. Each referendum had two options. The preliminary preferences of each student were recorded, and before the final (secret) ballot, the students participated in randomly assigned group discussions of about five students each with one moderator. Some – randomly assigned – moderators were neutral while others made limited (scripted) interventions in favor of one option or the other. The course writing and exam requirements were then set for the class by the majority.

To be clear, the experiment had the approval of the university Human Subjects Committee.[4] The balanced nature of our experiment helped to ensure that the final results of the deliberative process would not be impacted. As it turns out, majorities were so large that the final results were not altered by the experiment. Moreover, the vast majority of the students reported that they favored the opportunity to participate in the decision-making over the course requirements. No student expressed objection (not even in the anonymous evaluations).

Our experiment took place in a relatively tough environment for moderators to have an effect. First, unlike in community policy-making settings, our participants had well-defined a priori preferences over issues that they cared about and understood thoroughly (course requirements). This is often not the case in community settings. Second, moderators had no relationship with the students. We employed graduate students who had no authority over the students. We believe our moderators enjoyed no special status other than being called "moderators." In community settings, moderators are often community leaders or officials, who have some *direct* power over the participants, or NGO employees, who have a better understanding of the issues at stake and stronger

[4] The letter of exemption – under 45 CFR 46.101(b)(2) – is available on request.

communication skills than most of the participants. Finally, and most importantly, we deliberately constrained the ways in which our moderators could intervene in discussions (limiting them to three scripted statements). In community settings, such restrictions are absent.[5]

The results of our experiment are unambiguous: In the control group (90 cases), the overall rate of change between pre- and post-deliberation preferences was less than 17 percent. In cases where treatment moderators influenced *towards* a change in preference, the rate of change was over 37 percent (of 66 cases). In cases where treatment moderators influenced *against* a change in preference, not a single student changed opinion (0 out of 58). The results indicate that if a reasonable degree of disagreement existed in a community, organized interest groups could hijack a deliberative decision-making process if they had influence over the moderators.

We proceed by reviewing the experimental literature on deliberative democracy (section 3), and then describing our experiment in detail (section 3). Section 4 presents the results, while section 5 discusses the internal and external validity of the experiment. Section 6 concludes.

2. Experimental studies

Deliberative democracy has been employed at increasing rates around the world. The World Bank, the IMF, and the UN have all promoted its use. In both developing and developed countries NGOs use deliberative forums to decide local projects. In New Haven, Connecticut, for example, the Community Foundation for Greater New Haven used a deliberative forum to prioritize public

[5] From February to April 2009, one of the authors of this paper participated in the weekly meetings of the Participatory Budgeting Council in Porto Alegre, Brazil. The author observed moderators who were not neutral and used tactics including interruptions, shouting, strategic agenda setting, and even direct threats towards participants. For example, they warned that participants who criticized the city government would not receive city funding for their projects.

spending. And since 1994, there have been more than twenty-two Deliberative Polls® conducted in various countries. Researchers have proposed "Deliberation Day" as a pedagogical national event around election time in the United States, claiming it would transform electoral campaigns (Ackerman and Fishkin 2004).

Experiments have been conducted to test various facets of deliberative discussions, such as the preferences of participants.[6] Farrar et al. (2009a), for example, find that the effect of deliberation is smaller for salient issues, and when the participants are better informed (also see Luskin et al. 2007). Others show that deliberation makes the preferences of participants more structured and orderly (List et al. 2006).

Some studies have tested the effect of inequality among the participants on the deliberative outcome. Morrison and Singer (2007) find that inequality among the participants affects the perception of the outcomes of deliberation. Other studies have investigated the effect of group pressure on the outcome of group discussion. Farrar et al. (2006), List et al. (2006), and Luskin et al. (2007) find that participants modify their preferences to conform with their group.[7]

---

[6] We focus on the experimental literature, but note that non-experimental applications of participatory and deliberative democracy have been made in business, law, medicine, game theory, and political science. Our understanding of deliberative democracy has been particularly influenced by Ackerman (1991), Barber (1984), Benhabib (1992), Chambers (1996), Cohen (1997), Dryzek (2000), Fung (2004), Gaventa (2006), Gutmann and Thompson (1996), Fishkin and Luskin (2005), Habermas (1984), Nino (1996), Pateman (1970), Risse 2000, and Shapiro (2003). For reviews, see Simone (2003), Chambers (2003), and Mendelberg (2002).

[7] Their results contrast with the psychology literature that analyzes group polarization effects. Sophisticated research is emerging on strategic behavior in deliberations. In a non-experimental

A number of authors have argued that moderators might affect the deliberative process (for example, Sanders 1997, Young 2002, Shapiro 2003, Mutz 2008). We know of only two studies, however, that test the effect of moderators on deliberation. Fulwider (2005) finds that the presence of moderators makes opinion change more likely. Similarly, Humphreys et al. (2006) find that moderators have an effect. They analyze the results of a national deliberation organized by the UN Development Program in São Tomé and Príncipe, where citizens participated in moderated group discussions to decide country-wide economic priorities. The moderators were randomly selected from a group of civil society leaders and public officials, with some moderators randomly assigned to more than one group. The authors infer the importance of moderators from the share of the variance explained by moderator-specific effects. Moderator effects account for over one-third of variation in the outcomes. The authors also provide evidence that opinions changed in the direction of moderator influence. Moderator preferences – which were not manipulated by the experimenters – were recorded one week after the deliberation.[8] Moderators were assigned randomly, but their preferences were not (see Imai and Yamamoto 2008 for a discussion).

This prompts us to ask two questions. Can moderators influence participants to *change* their opinions? Can their influence also *reinforce* existing preferences? Our experiment is designed to answer these questions directly. It is the first study to randomly assign the direction of moderator

---

study, Muhlberger (2007) finds that participants are minimally strategic, but for contrasting views readers should see Landa and Meirowitz (2009), Dickson et al. (2008a,b), Hafer and Landa (2007), Meirowitz (2007), Hafer and Landa (2005), and Landa (2005). For a recent study that combines game-theoretic and psychological approaches tested with an experimental design, see Myers (2010).

[8] Moderators may have been influenced by the group (Humphreys et al. 2006:598), but this is unlikely as it implies strange moderators who are both highly influential and highly malleable.

influence. Moderator "preferences" are randomly assigned, and thus moderator influence is not correlated with any other observed and unobserved variables.

3. The experiment

The field experiment took place in an Introduction to Comparative Politics course with an enrollment of 107 students. There were no lectures or readings on deliberative democracy, but the students did learn about elections for the first month of the course. The course syllabus noted the following: "After studying elections, we will be holding elections in class ourselves! Students will be able to vote for alternative writing requirements and alternative exam requirements. Majority rules." The syllabus provided options for the writing and exam requirements (see table 1).

Table 1: Writing and exam options

	Writing requirement		Exam requirement
A)	3 short response papers (maximum 3 pages each)	A)	Cumulative final exam (45%).
B)	1 short response paper (maximum 2 pages) and 1 final paper (paper should be 8 pages)	B)	Midterm examination (covering lectures 1-15) and final examination (covering lectures 17, 19-25).

During their respective section-meetings, students were divided into groups of about five, and they discussed each requirement (writing and exam) for 15 minutes. Each group had a moderator who was introduced as a graduate student. No moderator was serving (or had ever served) as the teaching assistant for anyone in his/her group, so the moderators had no grading power. There were a total of 12 moderators (four from our class sections, eight from outside). They each moderated between two and five of the 24 total groups. Their "preferences" were randomly assigned. Note that the assignment of preference was randomized according to the group, not the specific moderator. Treatments for each discussion were independent, and all combinations of treatments A, B and placebo were possible between the writing and exam discussions.

As noted in the introduction, the graduate students were not viewed as holding authority over undergraduates. We believe that the only authority attributed to them was the implicit power any moderator has. Indeed, we suspected that this position alone might make them influential.

To test this hypothesis, we randomly assigned moderators to intervene in discussions by expressing opinions (or not) over the various options. We scripted the interventions and report exactly what the moderators said in the web appendix. The fact that some moderators would intervene in the discussions was unknown to the students ex ante.

"Treatment" moderators made three statements supporting their "preferred" option. The first statement was read at the beginning of the discussion; the second, halfway through the discussion; and the third at the end. Other than this, treatment moderators could only attempt to influence the discussion using body language, nodding their heads during the discussion when students raised points that agreed with their "preference," shaking their heads when students disagreed. "Placebo" moderators also read three statements at exactly the same time intervals as the treatment moderators, but these expressed no opinion about the options. They did not use body language, and they were instructed to be neutral.[9]

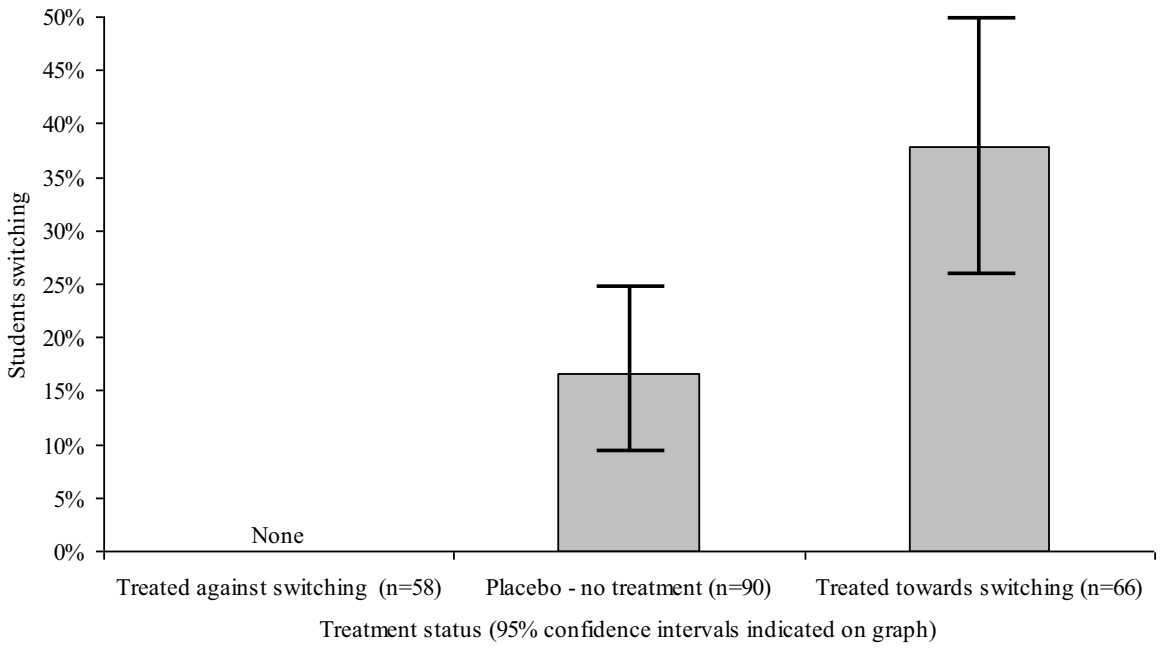
The students first discussed the writing requirement and then voted for their preferred option by secret ballot. Next the students discussed the exam requirement and then voted for their preferred option by secret ballot. After this, the students filled out a survey, which asked two groups of standard questions aimed to identify the participant and to evaluate her/his perceptions of the process.

[9] "Preferences" were randomly assigned as entire scripts per discussion. So moderator preference was consistent throughout a single discussion. See the appendix for details.

4. Results

As we turn to the results, it is noteworthy to mention the informal feedback we received from our moderators. Treatment moderators reported they felt paralyzed by the script and felt this limited them from having much influence. We decided in advance that scripting the interventions of the moderators would limit the effect of moderators who are particularly charismatic (or not). We also felt that scripting interventions would make clear the level of intervention, and make the experiment easy to explain and to replicate. Finally, we wanted to see if even minimal intervention could have an effect. It did. Figure 1 and Table 2 summarize the results.

Figure 1: Summary of Results: how many students switched preferences?



The mean for the *placebo* group is 17 percent, and a t-test indicates that we can say with 95 percent confidence that the hypothetical population mean is between 9 percent and 25 percent; we reject the null hypothesis of a mean of zero with greater than 99 percent confidence. This is important. It implies that in the absence of a treatment, we expect some people to change opinion – in other words, deliberation is consequential in the placebo group.

The mean for the group treated *against* change is 0 percent change. This result is so strong, it defies standard statistical techniques to test for uncertainty (the effect appears to be so strong that there is no variation in the dependent variable). For the group treated *towards* change, the mean is 38 percent, and a t-test indicates we can say with 95 percent confidence that the rate of change in the hypothetical population is between 26 percent and 50 percent.

Table 2 also shows the differences across the rates of change in these three groups are also highly statistically significant. According to a t-test comparing the means of the two samples (unpaired), within 99 percent confidence, the rate of change with a treatment *against* change is lower by between 4 and 30 percent compared to the placebo group. The rate of change with a treatment *towards* change is higher by between 3 and 39 percent compared to the placebo group, again within 99 percent confidence.

Table 2 Summary of results

Moderator treatment (# of students)	% changing opinion (# of students)	95% confidence interval
Treatment *against* changing opinion (n=58)	0% (0/58)	(0%, 0%)
Placebo - no treatment (n=90)	17% (15/90)	(9%, 25%)
Treatment *towards* changing opinion (n=66)	38% (25/66)	(26%, 50%)
Differences	**Absolute difference**	**99% confidence interval**
Treatment *against* changing vs. Placebo	17%	(4%, 30%)
Treatment *towards* changing vs. Placebo	21%	(3%, 39%)
Treatment *against* vs. Treatment *toward*	38%	(21%, 55%)

4.1 Descriptive data

We next present cross-tabulations of the data. Table 3 presents the detailed data – the percentages of students by their pre- and post-discussion preferences with the results from the writing requirement deliberation on top and the results from the exam requirement deliberation on bottom. Table 4 summarizes the results by experiment.

Consider first the writing requirement results. Generally, option A was preferred. Nonetheless, some students preferred option B, and there were students who switched. The switching trends consistently followed moderator influence:

- Among the students who started with a preference for option A, 100 percent of those treated for A stayed with it, 97 percent of those who received a placebo treatment stayed with it, and just 80 percent of those treated for B stayed with A.
- Among the students who started out preferring option B, 100 percent of those treated for B stayed with it, 63 percent who received a placebo treatment stayed with it, and just 25 percent of those treated for A stayed with B.
- Regarding indifference, there turned out to be no indifferent students who were treated for A. Among indifferent students who received the placebo treatment, half the students switched to A and half to B. Among the indifferent students treated for B, 100 percent switched to B.

Turning to the exam requirement, students generally preferred option B.

- Among students who started with a preference for option A, 100 percent of those treated for A stayed with it, 40 percent of students receiving the placebo stayed with it, and *none* of the students treated for B stayed with A.
- Among the students who started out preferring option B, 100 percent of those treated for B stayed with it, 100 percent who received a placebo treatment stayed with it, but only 79 percent of those treated for A stayed with B.
- For students who started out indifferent, those treated for A split evenly between A and B, those receiving the placebo split 17 percent to 83 percent in favor of option B, and 75 percent of the indifferent students treated for B switch to B with the remaining 25 percent remaining indifferent – no indifferent students treated for B switched to A.

All of the results are thus broadly consistent with our hypothesis about moderator influence.

Table 3a: The change in preferences pre-/post-deliberation by moderator treatment (%)

Writing requirement

Treatment for option A				Placebo				Treatment for option B			
	Post-discussion preference				Post-discussion preference				Post-discussion preference		
Pre-disc. pref.	A	B	Indiff.	Pre-disc. pref.	A	B	Indiff.	Pre-disc. pref.	A	B	Indiff.
A	**100%**	0%	0%	A	**97%**	3%	0%	A	**80%**	20%	0%
B	75%	**25%**	0%	B	38%	**63%**	0%	B	0%	**100%**	0%
Indiff.	0%	0%	0%	Indiff.	50%	50%	0%	Indiff.	0%	100%	0%

Exam requirement

Treatment for option A				Placebo				Treatment for option B			
	Post-discussion preference				Post-discussion preference				Post-discussion preference		
Pre-disc. pref.	A	B	Indiff.	Pre-disc. pref.	A	B	Indiff.	Pre-disc. pref.	A	B	Indiff.
A	**100%**	0%	0%	A	**40%**	60%	0%	A	**0%**	100%	0%
B	13%	**79%**	8%	B	0%	**100%**	0%	B	0%	**100%**	0%
Indiff.	50%	50%	0%	Indiff.	17%	83%	0%	Indiff.	0%	75%	25%

Table 3b: The change in preferences pre-/post-deliberation by moderator treatment (raw data)

Writing requirement

Treatment for option A				Placebo				Treatment for option B			
	Post-discussion preference				Post-discussion preference				Post-discussion preference		
Pre-disc. pref.	A	B	Indiff.	Pre-disc. pref.	A	B	Indiff.	Pre-disc. pref.	A	B	Indiff.
A	**30**	0	0	A	**30**	1	0	A	**20**	5	0
B	3	**1**	0	B	3	**5**	0	B	0	**4**	0
Indiff.	0	0	0	Indiff.	1	1	0	Indiff.	0	3	0

Exam requirement

Treatment for option A				Placebo				Treatment for option B			
	Post-discussion preference				Post-discussion preference				Post-discussion preference		
Pre-disc. pref.	A	B	Indiff.	Pre-disc. pref.	A	B	Indiff.	Pre-disc. pref.	A	B	Indiff.
A	**1**	0	0	A	**2**	3	0	A	**0**	4	0
B	3	**19**	2	B	0	**38**	0	B	0	**23**	0
Indiff.	1	1	0	Indiff.	1	5	0	Indiff.	0	3	1

Table 4 Summary of results by deliberation

Writing deliberation		
Moderator treatment (# of observations)	% changing opinion	95% confidence interval
Treatment *against* changing opinion (n=34)	0%	(0%, 0%)
Placebo - no treatment (n=14)	15%	(3%, 26%)
Treatment *towards* changing opinion (n=32)	34%	(17%, 52%)
Differences	Absolute difference	95% confidence interval
Treatment *against* changing & Placebo	15%	(2%, 27%)
Treatment *towards* changing & Placebo	20%	(0.1%, 39%)
Treatment *against* & Treatment *toward*	34%	(18%, 51%)
Exam deliberation		
Moderator treatment	% changing opinion	95% confidence interval
Treatment *against* changing opinion (n=24)	0%	(0%, 0%)
Placebo - no treatment (n=49)	18%	(7%, 30%)
Treatment *towards* changing opinion (n=34)	41%	(24%, 59%)
Differences	Absolute difference	95% confidence interval
Treatment *against* changing & Placebo	18%	(2%, 34%)
Treatment *towards* changing & Placebo	23%	(3%, 42%)
Treatment *against* & Treatment *toward*	41%	(21%, 62%)

4.2 More about the moderators

The work of Humphreys et al. (2006) suggests that certain characteristics of moderators may make them more or less influential. We did not randomize on moderator characteristics, and thus we do not wish to make strong conclusions about them. We do want to make sure that our main findings are not, by chance, driven by idiosyncrasies in the assignment of moderators.

Our set of moderators was diverse. In total, there were 12 moderators. Eight were women (ages 23 to 33), and four were men (ages 26 to 31). Five of the moderators were US born (three of them women), while the others came from Italy, Jamaica, Korea, Spain, Poland (two moderators), and Turkey. We also asked them to self-rate their moderator ability, based on their previous experience. We thus consider the potential impact of whether moderators are male or female, international or US-born, native speakers of English, section teaching assistants, younger or older (with the median

being the cut-point), and moderate or high ability. Inspection of our data reveals just two patterns with respect to moderator characteristics and influence. Table 5a presents these results, with the notable findings shaded.

First, preference change was more likely if the moderator was a man in the *placebo* group. There is, however, no statistically significant difference between the effectiveness of male and female *treatment* moderators. We prefer not to speculate as to why students were more likely to switch in the presence of a placebo male moderator than a placebo female moderator. We did not randomize with respect to this characteristic, and it is not the focus of our experiment. We focus instead on the fact that when acting as treatment moderators, moderator-gender played no role. Whatever differences men and women may have as moderators, they disappear in our experiment once both take on a proactive role of exerting influence. For more on the impact of gender on moderator effectiveness, we refer readers to Humphreys et al. (2006).

Next, we find that moderators who reported having had a high level of (self-rated) ability were indeed more effective in getting students to switch preferences. The rate of change for high ability moderators was 41 percent (95 percent confidence interval: 32 to 66 percent), while the rate for moderate ability moderators was just 24 percent (95 percent confidence interval: 8 to 41 percent). We return to this below.

Other than these two observations our inspection reveals no other patterns that deviate from the findings presented above. The differences in placebo and treatment effects across different types of moderators are not statistically significant. (To see this, note on Table 5a that the 95 percent confidence intervals for the differences across moderator effects all overlap zero.)

Having shown that there is no statistically significant difference between the treatment effects across various moderator types, we turn to testing whether the treatment by each type of moderator is, itself, statistically significantly different from the placebo effect for that type of moderator. Here,

our primary results hold, at least at the 10 percent level of confidence, for all but two types of moderators. The exceptions are male moderators treating against change and moderate ability moderators treating towards change.

Regarding the first exception, even though no one who was treated against change switched preferences, the effect for male moderators is not statistically significant. Table 5b shows that the difference between the placebo group rate of change and the treatment group is relatively large (–24 percent). But there are only seven observations of students treated against change with male moderators (see table 5a). Thus, this appears to simply a problem of small sample size.

The other exception is for moderate ability moderators treating towards change. The rate of change for this placebo moderator-cohort is 14 percent, while the rate of change for this moderator-cohort treating toward change is 24 percent (see table 5a). The difference (–10 percent) is in the expected direction, but the two-tailed p-value is 0.30, indicating that we cannot reject the null hypothesis.[10] We might conclude that the best way to safeguard against moderator manipulation is to choose inexperienced moderators. Yet, we hesitate to draw strong conclusions along these lines because we did not randomize with respect to self-rated moderator ability, and our experiment is not designed to directly test this possibility. Nor did we establish rigorous guidelines for our moderators to rate themselves. Given the data that we collected on this variable, however, we do find it interesting that lower moderator ability presents the weakest treatment effect. Of course, even the moderate ability moderators had a statistically significant impact when they treated *against* change (p=0.03).

Indeed, the statistical significance of the difference between treatment and placebo holds for all other types of moderators. For female moderators treating against change, and younger moderators

[10] Note that self-rated moderator ability is not strongly related to age: rho=0.12 (n=12).

treating towards change, the difference is significant at the 10 percent level. For all other types, the significance is at the 5 percent level or higher.

Table 5a: More on the moderators

	Female v. male moderators											
	Treatment against change				Placebo				Treatment toward change			
	% switching			95% C.I.	% switching			95% C.I.	% switching			95% C.I.
Female	0%	(0/	51)	(0%, 0%)	6%	(2/	36)	-(2%, 13%)	33%	(15/	45)	(19%, 48%)
Male	0%	(0/	7)	(0%, 0%)	24%	(13/	54)	(12%, 36%)	48%	(10/	21)	(24%, 71%)
Difference	0%			(0%, 0%)	-19%			-(34%, -3%)	-14%			-(40%, 11%)
	International v. US-born moderators											
	Treatment against change				Placebo				Treatment toward change			
	% switching			95% C.I.	% switching			95% C.I.	% switching			95% C.I.
Internat'l	0%	(0/	35)	(0%, 0%)	15%	(8/	52)	(5%, 26%)	34%	(12/	35)	(18%, 51%)
American	0%	(0/	23)	(0%, 0%)	18%	(7/	38)	(6%, 31%)	42%	(13/	31)	(24%, 60%)
Difference	0%			(0%, 0%)	-3%			-(19%, 13%)	-8%			-(32%, 17%)
	Non-native v. native speakers of English											
	Treatment against change				Placebo				Treatment toward change			
	% switching			95% C.I.	% switching			95% C.I.	% switching			95% C.I.
Non-native	0%	(0/	24)	(0%, 0%)	15%	(8/	52)	(5%, 26%)	36%	(10/	28)	(17%, 55%)
English	0%	(0/	34)	(0%, 0%)	18%	(7/	38)	(6%, 31%)	39%	(15/	38)	(23%, 56%)
Difference	0%			(0%, 0%)	-3%			-(19%, 13%)	-4%			-(28%, 21%)
	Non-teaching assistant v. teaching assistant moderators											
	Treatment against change				Placebo				Treatment toward change			
	% switching			95% C.I.	% switching			95% C.I.	% switching			95% C.I.
non-TA	0%	(0/	37)	(0%, 0%)			n/a		33%	(14/	43)	(18%, 47%)
TA	0%	(0/	21)	(0%, 0%)			n/a		48%	(11/	23)	(26%, 70%)
Difference	0%			(0%, 0%)			n/a		-15%			-(40%, 10%)
	Younger v. older moderators											
	Treatment against change				Placebo				Treatment toward change			
	% switching			95% C.I.	% switching			95% C.I.	% switching			95% C.I.
Younger	0%	(0/	35)	(0%, 0%)	17%	(5/	30)	(3%, 31%)	36%	(14/	39)	(20%, 52%)
Older	0%	(0/	23)	(0%, 0%)	17%	(10/	60)	(7%, 26%)	41%	(11/	27)	(21%, 61%)
Difference	0%			(0%, 0%)	0%			-(17%, 17%)	-5%			-(29%, 20%)
	Moderate v. high ability moderators											
	Treatment against change				Placebo				Treatment toward change			
	% switching			95% C.I.	% switching			95% C.I.	% switching			95% C.I.
Moderate	0%	(0/	31)	(0%, 0%)	14%	(5/	36)	(2%, 26%)	24%	(7/	29)	(8%, 41%)
High	0%	(0/	27)	(0%, 0%)	19%	(10/	54)	(8%, 29%)	49%	(18/	37)	(32%, 66%)
Difference	0%			(0%, 0%)	-5%			-(21%, 11%)	-25%			-(48%, -1%)

Table 5b: More on the moderators

	Treatment *against* changing vs. *Placebo*		Treatment *towards* changing vs. *Placebo*	
	Difference	2-tail p-value	Difference	2-tail p-value
Female	-6%	0.09	-28%	0.00
Male	**-24%**	**0.15**	-24%	0.05
International	-15%	0.01	-19%	0.04
US-born	-18%	0.03	-24%	0.03
Non-native English	-15%	0.04	-20%	0.04
Native English	-18%	0.01	-21%	0.04
Non-TA	-17%	0.01	-16%	0.04
TA			n/a	
Younger	-17%	0.01	-19%	0.08
Older	-17%	0.04	-24%	0.01
Moderate ability	-14%	0.03	**-10%**	**0.30**
High ability	-19%	0.02	-30%	0.00

4.3 Estimating the probability of switching

In this section, we put our evidence to additional tests, introducing a number of control variables: (1) a discussion rule indicator variable, (2) the number of *other* students in the discussion group who push towards changing preference (according to their pre-discussion preference), (3) an indicator of athlete-status, and (4) an indicator of student gender.[11]

Regarding the discussion rule indicator, we randomly assigned some groups to follow a strict participation rule where every student had to speak before someone could speak twice, and speaking time was limited to two minutes per student. Other groups had no guidelines other than the common courtesy of listening to one another. The variable is coded 1 for students participating in groups with the discussion rule and 0 otherwise. We introduced this variable following the example of

[11] Freedman (2008a,b) questions the use of regression analysis with experimental data, as randomization does not justify the model. He suggests "cross-tabulation before regression," arguing that parametric analysis be used only to validate rates and averages for the treatment and control groups. This is our intention. Further note that Green (2009) shows the use of pre-treatment covariates in samples of n>20 can generate gains in precision without adding substantial bias.

participatory debates in the city of Porto Alegre where in some assemblies a strict three minute rule is enforced to allow everybody to have a chance to talk and to prevent the discussion being monopolized by some participants (Wampler 2007).[12]

We also constructed a variable measuring the number of *other* students in the discussion group who push towards changing preference. Note that for each observation of a student, this variable is constructed by counting the number of students in each group who push towards an option minus the student in question. This variable is inspired by the early work of Asch (1948) and the recent work of Sunstein (2000), who finds that group pressure can have a strong effect, and Luskin et al. (2007) and Farrar et al. (2006) who find little evidence of group effects. Although we do not seek to take sides on this debate, we introduce the variable both as a control and to compare the effect of a moderator versus the effect of simply any other additional group member advocating a position.

We introduce the athlete indicator, coded 1 for students involved in varsity sports, and 0 otherwise, because of an obvious correlation between athlete status and section. Most sections met in the afternoon, when athletes had practice. So athletes tended to enroll in one of the two morning sections available. Furthermore, we suspected that athletes might have stronger preferences on assignments than other students, since they have particularly strict schedules and certain assignment dates might suit them better than others.

Finally, we also control for student gender.[13]

---

[12] See Mutz (2008:530) for an interesting discussion on the potential role of moderators along these lines. For substantive work on the subject, see Dickson, Hafer, and Landa (2008b).

[13] On the question of gender, we also interacted the gender of the moderator and that of the participants, but found no robustly significant effects. Humphreys et al. (2006) also consider participant age, but with our sample of undergraduates, there is not much variance.

We model the probability that a student *changes* his or her preference, analyzing four *transition* probabilities: going from option A to B or B to A for each of the two discussions. In all four cases, the dependent variable is conditioned on a pre-discussion preference against that option. Thus, if the dependent variable is coded 1, it indicates a "transition" or change in preference toward the option in question; if it is coded 0, it indicates the student did not switch to the option in question. This is a high standard – there are some students who moved from one option towards indifference, following moderator influence, but we do not count these. We do, however, allow pre-discussion indifferent students to switch in either direction. As the dependent variable is dichotomous, we employ a logit model. Importantly, we cluster errors by group because randomization (of moderator preference) is at the group level.[14]

Table 6a presents the results for all four of the transition probabilities described above. In each case, we begin with a "barebones" model, introducing only the treatment variables of interest: moderator treatment towards option A or B. Note that in three of our four models, we cannot introduce the treatment variable against switching because it predicts the dependent variable *perfectly*. Recall from above, no one who received treatment B switched to A for papers or exams, and no one who received paper treatment A switched to option B. These results are so perfectly in line with our hypothesis that they defy estimation using a logit model. Indeed, only one student subjected to exam treatment A actually switched to B, and this student started out as indifferent. After the barebones model, we introduce the control variables mentioned above.

Taking the results in turn, paper treatment A has a positive effect on switching to option A. But with only 21 observations, the result is not statistically significant at conventional levels. None of the control variables have statistically significant effects either. Of note, the percentage of others in

[14] We also clustered errors by moderator – all of the results reported below hold when we do so.

the group favoring option A has a negative effect on switching to A, although not a statistically significant one.[15]

Paper treatment B similarly has a positive effect on switching to paper option B, and the result is statistically significant at the 0.05 percent level. The result holds (and gets stronger) when we introduce control variables, which also have interesting effects. The discussion rule has a positive effect significant at the 0.05 percent level, indicating that students in groups where everyone was required to speak in turn with time-limits are more likely to switch. The group effect here is also quite strong. The percentage of others favoring paper option B has a positive effect on switching to B, significant at the 0.01 percent level. We were curious about comparing the group effect to the moderator effect, so we used Clarify software to simulate the impact of introducing a moderator versus increasing the percentage of others supporting option B from 0 to 25 then to 50 then to 75 and finally to 100. We find the mean effect of the moderator to be larger than the mean group effect up to 50 percent, and the mean group effect to be larger beyond that. But the confidence intervals overlap throughout (results available on request). Finally, we find that men are more likely to switch than women in this case, though this finding holds only in for switching to paper option B and is significant at the 0.10 percent level.

Exam treatment A also has a positive effect on switching to exam option A, significant at the 0.05 level. The result holds when we introduce the control variables, none of which are statistically significant. The percentage of others in the group favoring A has a positive effect on switching to A, but not a statistically significant one.

Next we turn to the probability of switching to exam option B. This was the only case where a student went against moderator influence, so we can estimate the impact of both treatments using

[15] In some of the other model specifications we tried, the treatment effect was statistically significant, although this effect is not robust.

the logit model. Exam treatment A has the expected negative effect on switching to B, and exam treatment B has the expected positive effect on switching to B, but neither appears statistically significant. Once we introduce control variables, however, the statistical significance of the effect of treatment A obtains at the 0.001 level. The effect of the discussion rule also has a significant effect, although in the opposite direction as reported above – students in groups where everyone was required to speak in turn with time-limits are less likely to switch. We must drop the male student indicator variable because it predicts the dependent variable perfectly – only unlike the statistically significant effect reported for this variable above, this time it is women who switched (all four observations of women switched to exam option B).

Summarizing, our hypothesis of moderator influence is confirmed throughout eight tests:

- The effect of paper treatment A on switching to A: positive but significant only at the 0.15 level.
- The effect of paper treatment B on switching to A: negative and predicts perfectly.
- The effect of paper treatment A on switching to B: negative and predicts perfectly.
- The effect of paper treatment B on switching to B: positive and significant at the 0.05 level.
- The effect of exam treatment A on switching to A: positive and significant at the 0.05 level.
- The effect of exam treatment B on switching to A: negative and predicts perfectly.
- The effect of exam treatment A on switching to B: negative and significant at the 0.001 level.
- The effect of exam treatment B on switching to B: positive but not significant.

We hesitate to draw strong conclusions about our pre-treatment control variables. We simply point out that the effect of moderator treatment survives their inclusion and is robust throughout our tests. Interestingly, while we find that the moderator effect is robustly statistically significant, the effect of other people in the group advocating a particular position is not, except in one setting, where the coefficient is smaller than the coefficient for the moderator. In all of the regressions where the group effect is not statistically significant, we can also say with 90 percent confidence

that the difference between its effect and the effect of the moderator is statistically significant. In the regression where the coefficient is statistically significant, we can say with 85 percent confidence that the difference between its effect and the effect of the moderator is statistically significant. This suggests that the moderator is not just another voice in the crowd, but has inherent influence.[16]

Finally, we return to the moderator-specific attributes that appeared to be significant in section 4.2 and introduce them as control variables.[17] See Table 6b. All of the main qualitative findings from Table 6a hold, except for switching to exam option B. Here, treatment against change is not statistically significant, but treatment towards change is (in table 6a, the reverse is true).[18]

[16] It would be interesting to randomly assign treatment "participants," who advocate an assigned "preference." This would enable researchers to directly distinguish between moderator and participant effects. Unfortunately, we did not have close enough working relationships with sufficient undergraduate participants since secrecy was of importance.

[17] We lose too many degrees of freedom when we include all control variables in the samples with only about 20 observations; in the regressions with more observations, the main findings hold with additional control variables we tried. These results are available on request.

[18] The number of observations is smaller because some of the moderator variables predict perfectly. Observations, along with these variables, are dropped from the estimation.

Table 6a: Estimating transitions (logit model)

	Switching to paper option A		Switching to paper option B		Switching to exam option A		Switching to exam option B	
Variable	Barebones	Control variables	Barebones	Control variables	Barebones	Control variables	Barebones	Control variables
Moderator treatment toward option A	2.28 (1.52)	3.63 (3.26)	^^	^^	2.54** (1.23)	3.65*** (1.21)	-1.67 (1.23)	-20.57*** (2.83)
Moderator treatment toward option B	^	^	2.50** (1.20)	6.25** (2.70)	^^^	^^^	0.97 (1.32)	-0.07 (1.41)
Discussion rules indicator		-2.36 (1.52)		4.63** (1.96)		0.74 (1.31)		-17.62*** (1.70)
# of others in group favoring option A		-0.93 (1.23)				1.64 (1.22)		
# of others in group favoring option B				3.78** (1.57)				1.43** (0.70)
Athlete indicator		0.06 (1.38)		1.65 (1.19)		-0.10 (1.15)		0.54 (1.34)
Male student indicator		-0.49 (0.88)		4.28* (2.45)		1.04 (0.72)		^^^^
Constant	-1.18 (0.93)	1.65 (2.16)	-3.42*** (1.05)	-16.49** (7.28)	-4.25*** (1.02)	-6.69** (2.92)	0.98 (0.99)	15.22 .
Number of observations	21	21	91	91	97	97	22	22
Log pseudolikelihood	-11.52	-8.74	-25.62	-9.60	-16.42	-14.99	-11.37	-7.73

^ Moderator treatment toward option B is dropped because the variable predicts the dependent variable perfectly. No one who was treated for paper option B switched to paper option A.

^^ Moderator treatment toward option A is dropped because the variable predicts the dependent variable perfectly. No one who was treated for paper option A switched to paper option B.

^^^ Moderator treatment toward option B is dropped because the variable predicts the dependent variable perfectly. No one who was treated for exam option B switched to exam option A.

^^^^ Male student indicator dropped because it predicts the dependent variable perfectly. The four observations of women all switched to exam option B. *** significant at the 0.01 percent level. ** significant at the 0.05 percent level. * significant at the 0.10 percent level. Robust standard errors, clustered at the group level, are in parentheses.

Table 6b: Estimating transitions (logit model) with moderator control variables

Variable	Switching to paper option A	Switching to paper option B	Switching to exam option A	Switching to exam option B
Moderator treatment toward option A	-0.51 .	^^	2.70** (1.30)	0.00 (0.89)
Moderator treatment toward option B	^	4.49** (1.52)	^^^	15.90** (0.81)
Male moderator	-13.29 .	1.28 (1.02)	-0.80 (1.60)	^^^^
Moderator ability (high/moderate)	61.59 .	2.31** (0.99)	^^^^	^^^^
Moderator age (23-31)	-12.12 .	0.21 (0.25)	0.41 (0.27)	5.07** (0.03)
Constant	309.19 .	-12.26 (7.50)	-15.14* (7.87)	-157.10 .
Number of observations	21	91	54	12
Log pseudolikelihood	0	-19.68	-13.84	-7.45

^ Moderator treatment toward option B is dropped because the variable predicts the dependent variable perfectly. No one who was treated for paper option B switched to paper option A.
^^ Moderator treatment toward option A is dropped because the variable predicts the dependent variable perfectly. No one who was treated for paper option A switched to paper option B.
^^^ Moderator treatment toward option B is dropped because the variable predicts the dependent variable perfectly. No one who was treated for exam option B switched to exam option A.
^^^^ Dropped because of estimability. Also note that the number of observations differ from table 6a because some observations are completely determined and dropped.

*** significant at the 0.01 percent level. ** significant at the 0.05 percent level. * significant at the 0.10 percent level. Robust standard errors, clustered at the group level, are in parentheses.

4.4 Discussion

Three findings strike us as particularly noteworthy.

(1) Moderators have an effect. This confirms the findings of Fulwider (2005) and Humphreys et al. (2006) in an original setting, where moderator preference is randomly assigned.

(2) In both discussions, one option was overwhelming preferred to the other in both pre- and post-preferences. Changes in preference, where they did occur, tended towards these majority preferences.

So for the writing requirement, 86 students started out in favor of option A (three short papers), while only 16 students started out in favor of option B (one long, one short paper) with five indifferent students. Switching preferences, when it occurred, was more likely to go towards

option A than towards option B. Seven out of 21 students (33 percent) switched to A, but only 10 out of 91 students (11 percent) switched to B. The results in Table 6 indicate that the moderator treatment effect is stronger and more statistically significant for switching to the unpopular option B. (The other finding – switching to the popular option A – is likely weaker simply because the sample size is obviously reduced when considering likelihood of the minority switching to the majority.)

For the exam requirement we see a similar pattern. Here, the overwhelming preference was for option B (two exams): 85 students started out in favor of this. Only 10 students started out in favor of option A (one cumulative exam), and 12 were indifferent. As for switching, a noteworthy 16 out of 22 students (73 percent) switched to option B, with only five out of 97 students (5 percent) switching to option A and two out of 95 students (2 percent) switching to indifference (though both of them switched away from option B). Again the results in Table 6 indicate that the moderator treatment effect is more robust for switching to the unpopular option (in this case option A). (Again, the effect of switching to the popular option B is likely weaker simply because the sample size is reduced when considering likelihood of the minority switching to the majority.)[19]

(3) Moderator effects were considerably stronger for reinforcing opinions than for changing them. Our most striking result – holding for one hundred percent of cases – is that deliberation is unlikely to lead someone to switch preferences if he or she is in agreement with the moderator to begin with. In this sense, moderators can undermine the very point of deliberation.

[19] Future research should consider the interaction between moderator influence and (minority/majority) group opinion.

5. Validity

In this section, we consider some questions regarding the external and internal validity of our experiment. Concerning external validity, the environment of our experiment is very different from real-life applications of deliberative decision-making processes. Yet we believe that the differences may actually bias the results *against* the moderators influence that we observed.

In most real-life applications of deliberative decision-making, the results of the process are an allocation of funds to a project. Most of the time, this project affects the life of only a subset of the participants in the deliberative process (for example, the construction of a park in a specific district of the city, or the allocation of funds to a school or to a public mill). So the level of issue-salience may vary a great deal from individual to individual, with some participants directly affected by the policy under deliberation, and others only indirectly or not at all. In our experiment, the outcome directly affected the life of all of the participants regarding an activity that shapes the biggest part of Ivy-league life – their course requirements. Therefore, we believe the salience of the issue in our experiment to be relatively high for each individual taking part in the deliberation. Experimental work in this area indicates that deliberation has less impact on highly salient issues (see, for example, Farrar et al. 2007). Presumably people are less easily swayed because their preferences are strong.

With regard to the subjects of our field experiment, while not a representative sample of the population, we suspect that attending college and selecting an Introduction to Comparative Politics Class, on average, has little to do with the predisposition of the students to be influenced. Yet, even if students tend to change opinion more (or less) than the rest of the population, there is a problem of external validity only if, furthermore, students are more prone to the effect of moderators. We find this particular combination of conditions to be unlikely. Ultimately, of

course, this is an empirical question that we cannot answer until we replicate the design outside the university setting using a random sample of citizens.

With regard to our moderators, graduate students at our university do not have special status with respect undergraduates, and the scripts limited their influence. In real life, moderators often enjoy special status and can intervene much more effectively than our actors were allowed to do.

Still, some questions may remain. Notably, students clearly had overwhelming preferences for three papers over two and two exams over one, something we did not anticipate. Both the pre- and post-deliberation preferences – as well as the effects of moderators – reflect this. This resonates with recent research showing that deliberation is more likely to induce change when participants have less well formed preferences (Farrar et al. 2009a, Luskin et al. 2007). It also may indicate that students who switched from the less popular options to the more popular options did not care as much. Thus, the magnitude of the moderator effect may depend on the issue and on the strength of preferences. But it does not invalidate our qualitative conclusions regarding differences between the treatment and the control groups, and even the lower bounds on our estimated effects of moderators on inducing and preventing change are greater than zero.

Moderator ability is a related question. As explained above, the design of the experiment forced moderators to intervene minimally and to follow a precise script. Still, there could be some residual effect of personal charisma. Treatments were assigned randomly to moderators, but it is possible that all the "high ability" moderators were – by chance – assigned to just one treatment in the discussion. Yet, again, this cannot explain the difference between the treatment and the control groups. In the control group, the ability of the moderator plays no role. Another related problem could be the potential differences in the quality of the scripts we adopted, but a similar logic applies. It is possible that some of the scripted arguments are more convincing than

others. Again, this would only impact the difference between the two treatment effects, not the difference between the treatment group and the control group.

A further problem, related to the *internal* validity of the experiment, regards the distribution of the students across sections and days (Tuesday, Wednesday, Thursday, and Friday). The students were not randomly assigned to a section or day, although they were randomly assigned to a group and to a treatment within their section. A student's section-choice is usually quasi-random and is typically linked to the schedule, so clusters of similar students might end up in the same section. We tested for this type of clustering and found that this significantly increased the standard error for only one of our findings – the effect of paper treatment A on switching to A – which we already reported above not to be statistically significant at conventional levels. All of our other results are robust to clustering standard errors by day, section, group, or moderator.

6. Conclusion

Communities around the world are increasingly turning to deliberative and participatory processes where moderators engage people in direct policy-making, and the approach has been advocated by the World Bank, the IMF, the UN, NGOs, and municipal governments. The results of our study indicate that this approach to decision-making is at risk of manipulation.

Our objective is not to criticize the normative or the experimental literatures on deliberative and participatory approaches, where moderators are neutral by design. Rather, we seek to raise a flag: when the deliberative approach enters into the realm of real politics, interest groups try to manipulate it. Our goal has been to simulate one way an interest group might try to exercise influence: through the moderators.

Our field experiment was *not* tailored to affect the overall decision-making process. We randomly assigned preferences to treatment moderators that were balanced: half of them

influenced for one option, half for the other. There were also placebo moderators who exercised no influence. Moreover, we scripted limited interventions for our moderators, and our subjects were well-informed with strong preferences over the issues at stake. Nevertheless, the results of moderator influence are unambiguous. Regarding the ability of moderators to induce a change in opinion, our aggregate data indicate a 21 percent increase in the rate of change (with a 95 percent confidence interval of 3 to 39 percent) over the placebo group rate of change. Regarding the ability of moderators to prevent a change in opinion, our aggregate data indicate a 17 percent decrease in the rate of change (with a 95 percent confidence interval of 4 to 30 percent) over the placebo group rate of change. If enough disagreement exists over an issue, such that a community is divided evenly, a three to four percent shift in opinion could be pivotal. So, even the lower bounds of our estimates indicate grounds for concern. Happily, we can report that our experiment had no effect on the ultimate group decision because of the way we balanced the treatments.

But imagine if a well-organized interest group were to manipulate moderators to influence group discussions towards a specific outcome in an unbalanced manner. Our results suggest that such an endeavor could be fruitful. The interest group might be able to sway the outcome of the deliberative process on an issue where a reasonable degree of disagreement and uncertainty exists. We conclude that deliberative democracy is at risk if no one moderates the moderators. If the selection of even just some of the moderators is left in the hands of special interest groups, they might be able to sway the result of a deliberative decision. Thus we suggests that the practitioners and participants be made aware of this potential problem. Moderating the

moderators should be a priority for any organization interested in deliberative decision-making processes – for interest groups are clever and will start with them.[20]

References

Abers, R. 1998. From Clientelism to Cooperation: Local Government, Participatory Policy, and Civic Organizing in Porto Alegre, Brazil. *Politics and Society* 26:511-37.

Asch, S. E. 1948. The doctrine of suggestion, prestige, and imitation in social psychology. *Psychological Review* 41:258-290.

Ackerman B. 1991. *We the People: Foundations*. Cambridge, MA: Harvard Univ. Press.

Ackerman B., and J. Fishkin. 2004. Righting the Ship of Democracy. *Legal Affairs* (Jan-Feb):34-39.

Avritzer L. 2002. *New Public Spheres in Brazil*. New Jersey: Princeton University Press.

Baierle, S. G. 2007. *Lutas Urbanas em Porto Alegre: entre a revolução política e o transformismo*. Porto Alegre: Cidade – Centro de Assessoria e Estudos Urbanos.

Barber, B. R. 1984. *Strong Democracy*. Berkeley: University of California Press.

Benhabib Sheila. 1992. *Situating the Self*. New York: Routledge.

Chambers S. 1996. *Reasonable Democracy*. Ithaca, NY: Cornell Univ. Press.

Chambers, S. 2003. Deliberative Democratic Theory. *Annual Review of Political Science* 6:307-26.

Cohen, J. 1997. Deliberation and democratic legitimacy. In J. F. Bohman & W. Rehg (Eds.), *Deliberative democracy*. Pp. 67-91. Cambridge, MA: MIT Press.

Dryzek, J. K. 2000. *Deliberative Democracy and Beyond*. Oxford: Oxford University Press.

[20] Here we liberally paraphrase from Decimus Iunius Iuvenalis's *Satire VI*.

Dickson, Eric, Catherine Hafer, and Dimitri Landa. 2008a. Cognition and Strategy: A Deliberation Experiment. *Journal of Politics* 70(4): 974–89.

Dickson, Eric, Catherine Hafer, and Dimitri Landa. 2008b. Rules of Debate: Theory and Experiment. Typescript. New York University.

Farrar, Cynthia, Donald P. Green, Jennifer E. Green, David Nickerson, and Steven Shewfelt. 2006. Does Discussion Group Composition Affect Policy Preferences? Ms. Yale University.

Farrar, Cynthia, James S. Fishkin, Donald P. Green, Christian List, Robert C. Luskin, and Elizabeth Levy Paluck. 2007. Disaggregating Deliberation's Effects. Ms. Yale University.

Fearon, James, Macartan Humphreys, and Jeremy M. Weinstein. 2009. Can Development Aid Contribute to Social Cohesion After Civil War? Evidence from a Field Experiment in Post-Conflict Liberia. American Economic Association Annual Meetings, January, San Francisco.

Fishkin, J. and Robert C. Luskin. 2005. Experimenting with a Democratic Ideal: Deliberative Polling and Public Opinion. *Acta Politica* 40: 284-298.

Freedman, David A.. 2008a. Randomization Does Not Justify Logistic Regression. *Statistical Science* 23: 237–49.

—— 2008b. On regression adjustments in experiments with several treatments. *Annals of Applied Statistics* 2: 176–96.

Fulwider, J. 2005. Do Moderators Matter? Answering a Jury Deliberation Challenge to Deliberative Democracy. Paper presented at the Annual Meeting of the American Political Science Association, Marriott Wardman Park, Omni Shoreham, Washington Hilton, Washington, DC.

Fung, Archon. 2004. *Empowered Participation*. Princeton, NJ: Princeton University Press.

Gaventa, J. 2006. Deepening the Deepening Democracy Debate. IDS Working Papers – 264.

Gerber, Alan S., and Donald P. Green. 2008. Field Experiments and Natural Experiments. In Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier (eds.), *Oxford Handbook of Political Methodology*. New York: Oxford

Green, Donald P. 2009. Regression Adjustments to Experimental Data: Do David Freedman's Concerns Apply to Political Science? Paper presented at the 26th annual meeting of the Society for Political Methodology, Yale University, July 23-25.

Gutmann A, and D. Thompson. 1996. *Democracy and Disagreement*. Cambridge, MA: Harvard University Press.

Habermas, J. 1984. *Theory of Communicative Action, vol. 1: Reason and the Rationalization of Society*. (translated by Thomas McCarthy). Boston: Beacon Press.

Hafer, Catherine, and Dimitri Landa. 2005. Deliberation and Social Polarization. Ms. New York University.

Hafer, Catherine, and Dimitri Landa. 2007. Deliberation as Self-Discovery. *Journal of Theoretical Politics* 19(3): 329–60.

Humphreys, Macartan, William Masters, and Martin Sandbu. 2006. The Role of Leaders in Democratic Deliberations: Results from a Field Experiment in São Tomé and Príncipe. *World Politics* 58: 583-622.

Imai, K., and T. Yamamoto. 2008. Causal Inference with Differential Measurement Error: Nonparametric Identification and Sensitivity Analyses of a Field Experiment on Democratic Deliberations. Working Paper, Princeton University.

Landa, Dimitri. 2005. The Moral Economy of Deliberative Participation. Ms. New York University.

Landa, Dimitri and Adam Meirowitz. 2009. Game Theory, Information, and Deliberative Democracy. *American Journal of Political Science* 53 (2):427–444.

List, Christian, Robert C. Luskin, James S. Fishkin, and Iain McLean. 2006. Deliberation, Single-Peakedness, and the Possibility of Meaningful Democracy: Evidence from Deliberative Polls. PSPE Working Paper No. 1.

Luskin, Robert C., James S. Fishkin, and Kyu S. Hahn. 2007. Deliberation and Net Attitude Change. Paper presented at the ECPR General Conference, Pisa, Italy.

Mansuri, Ghazala and Vijayendra Rao. 2004. Community-Based and -Driven Development: A Critical Review. *The World Bank Research Observer* 19 (1):1-39.

Marquetti, A. 2003. Participação e redistribuição: o orçamento participativo em Porto Alegre. In: Avritzer, L.; Navarro, Z. (Org.). *A inovação democrática no Brasil*. Pp. 129-156. São Paulo, Cortez Editora.

Marquetti, Adalmir A., and Duilio de Avila Bêrni. 2006. Democracia Participativa, Performance Fiscal e Distribuição: a evidência dos municípios gaúchos. In: *Terceiro Encontro de Economia Gaúcha*. Porto Alegre.

Meirowitz, Adam. 2007. In Defense of Exclusionary Deliberation: Communication and Voting with Private Beliefs and Values. *Journal of Theoretical Politics* 19(3): 301–27.

Mendelberg, Tali. 2002. The deliberative citizen: Theory and evidence. *Political Decision Making, Deliberation and Participation* 6: 151–193.

Morrison, K. and Singer, M. 2007. Inequality and Deliberative Development: Revisiting Bolivia Experience with the PRSP. *Development Policy Review* 25 (6): 721-740.

Muhlberger, P. 2007 Strategic and Communicative Rationality in a Deliberative Field Experiment. Paper presented at the annual meeting of the Midwest Political Science Association, Palmer House Hotel, Chicago, IL.

Mutz, Diana C. 2008. Is Deliberative Democracy a Falsifiable Theory? *Annual Review of Political Science* 11:521–38.

Myers, C. D. 2010. Information Sharing in Small Group Deliberation. Ms. Princeton University.

Nino, C. S. 1996. *The Constitution of Deliberative Democracy*. New Haven: Yale University Press.

O'Brennan, J. 2004. Ireland's National Forum on Europe: Elite Deliberation Meets Popular Participation. *Journal of European Integration* 26 (2): 171-189.

Pateman, C. 1970. *Participation and Democratic Theory*. Cambridge University Press.

Platteau, Jean-Philippe and Frédéric Gaspart. 2003. The Risk of Resource Misappropriation in Community-Driven Development. *World Development* 31 (10):1687-1703.

Risse, Thomas. 2000. "Let's Argue!": Communicative Action in World Politics. *International Organization* 54 (1):1–39.

Shapiro, I. 2003. *The State of Democratic Theory*. Princeton University Press.

Sanders, L. M. 1997. Against deliberation. *Political Theory* 25, 3, 347–376.

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

Wampler, B. 2007. *Participatory Budgeting in Brazil: Contestation, Cooperation, and Accountability*. Pennsylvania State University Press.

Winters, Matthew S. 2007. Explaining Selection into Indonesia's Kecamatan Development Program. Paper Prepared for Presentation at the Annual Meeting of the Midwest Political Science Association.

World Bank. 2008. *Toward a More Inclusive and Effective Participatory Budget in Porto Alegre Vol I.* Report No. 40144-BR. Washington DC: World Bank.

Young, I. M. 2002. *Inclusion and Democracy*. Oxford University Press, Oxford.

APPENDIX

In this Appendix, we present the precise scripts used by our moderators in our experimental setting. Students in a political science class voted on options over their writing and exam requirements. Before voting, 107 students participated in group discussions of about five students each with one moderator. Some (randomly assigned) moderators remained neutral throughout the discussions, while others made limited interventions, supporting a specific option. The appendix also presents the survey used to record the students' pre-deliberation preferences

Detailed description of the experiment

The experiment proceeded as follows. After breaking into groups, the students were given a survey. We began by soliciting pre-discussion preferences, as shown in Table 1A. The second page of the survey read, "Please do not turn the page until the moderator instructs you to do so. Thank you."

Table 1A: Pre-discussion survey

ON THIS PAGE ARE PRE-DISCUSSION SURVEY QUESTIONS. ***THEY DO *NOT* COUNT FOR THE FINAL VOTE***. YOU WILL BE ALLOWED TO CHANGE YOUR MIND DURING THE DISCUSSION.
Select your preferred writing requirement option with an x:
Three short papers (maximum 3 pages each)
One short paper (maximum 2 pages) and one long paper (8 pages)
I am indifferent
Select your preferred exam requirement with an x:
One cumulative final
One midterm and one non cumulative final
I am indifferent

The discussion over the writing requirement ensued. Two different discussion rules were randomly assigned to groups (irrespective of moderator "preference" assignment). According to one rule, discussion was left completely open. Specifically, students were instructed by the moderator, "*All*

students can speak as they like. Please be polite and listen to one another." According to the second rule, discussion was regulated. Specifically, students were instructed, "*Each student must be given a chance to speak before anyone can speak a second time. To ensure that everyone gets a chance, a student can talk for a maximum of 2 minutes for his/her first turn, and 1 minute for his/her second and third turn. This rule will be strictly enforced and I apologize in advance if I have to cut anyone off. But I will if necessary. I thank you in advance for your cooperation and understanding.*"

Next, moderators read one of the scripts from Table 2A, either expressing a preference or not. The students were then allowed to discuss the writing requirement. When the discussion time reached the halfway point, the moderators continued with the scripts in Table 3A. Student discussion was allowed to continue until the time limit was reached. When time was up, the three moderators concluded with the final intervention described in Table 4A. Students then voted on the writing requirement by secret ballot. They could choose between the two options above or choose indifference. After voting, the deliberation turned to the exam requirement. The experiment for this discussion followed the exact same format as the writing requirement discussion. "Preferences" of the moderators were randomly assigned (ahead of time).[21] The discussion rule that had been randomly assigned for the writing requirement deliberation was maintained for the exam requirement. We omit the scripts for the second discussion; they are available on request.

[21] Thus, some placebo moderators from the first discussion were randomly assigned to be influencing moderators for the second discussion (and vice versa). We instructed such moderators to say s/he did not have any preference over the second decision if asked. No one asked.

Table 2A: First discussion, opening intervention

Type of moderator	Script
Placebo	*As the TA has said, we have to deliberate first on the decision regarding the writing requirement. You have to decide between 3 short papers, 3 pages each, and 1 short paper and 1 long paper, 2 pages and 8 pages.*
In favor of three short papers	*Before letting you discuss, I want to share with you one little tip from my experience regarding this type of assignment. I take my job as moderator seriously and have notes: Don't be fooled by the fact that three assignments are more than two. Take into consideration the fact that writing a short paper might take less time than writing a long one. You might be able to write two good 3-page papers in less time than an 8-page paper takes.*
In favor of one long paper	*Before letting you discuss, I want to share with you one little tip from my experience regarding this type of assignment. I take my job as moderator seriously and have notes: Don't be fooled by the fact that writing a short paper requires less time than writing a long one. Writing three different short papers requires coming up with three different ideas. Coming up with ideas is the hardest part of this assignment. If you do one short and one long paper, Prof. Freeland will allow you to use the short paper as a draft for the long paper. This means that in the end, if you choose the option of one short and one long paper, you simply have to come up with ideas for one paper and then just write one longer paper using the same ideas.*

Table 3A: First discussion, midway intervention

Type of moderator	Script
Placebo	*Nothing*
In favor of three short papers	*At this point of the discussion I want to focus your attention on one important element: the preparation for the exam. Writing three short papers on three different topics means that you have already thoroughly prepared three major topics that will be covered in exam questions. This will give you more mastery of the subject and might help make studying for exam questions easier, regardless of the structure of exams.*
In favor of one long paper	*At this point of the discussion I want to focus your attention on one important element: the preparation for the exam. Writing three short superficial papers on three different topics means that you haven't mastered any single topic. On the other hand writing a long paper allows you to focus and master a topic. Thus, when the time of the exam comes, you have already prepared in depth some of the material for the final exam. This will give you more mastery of the subject.*

Table 4A: First discussion, final Intervention

Type of moderator	Script
Placebo	*The time is almost up. It's time to stop the discussion. You have raised excellent points to support both proposals.*
In favor of three short papers	*Before voting I want to (point out)/(remind you of)*[22] *one final element: the importance of learning parsimony. When you compete for grants or make presentations, you will discover that often you are allowed to write only a couple of pages. Thus, from the point of view of pure usefulness for your future, the 3 short page papers is probably best.*
In favor of one long paper	*Before voting I want to (point out)/(remind you of) one final element: the importance of learning how to write a draft and then learning how to expand this draft into a proper project. This will be the process you will follow to write your senior essay and any type of other project in the future. Thus, from the point of view of pure usefulness for your future, one short paper and one long paper is probably best.*

[22] Moderators were allowed to say "remind you" instead of "point out" if this point was already raised by students. We did not want the scripts to seem awkward.