

Does Enclave Deliberation Polarize Opinions?

Kimmo Grönlund · Kaisa Herne · Maija Setälä

Published online: 8 February 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract When like-minded people discuss with each other, i.e. engage in ‘enclave deliberation’, their opinions tend to become more extreme. This is called group polarization. A population-based experiment with a pre-test post-test design was conducted to analyze whether the norms and procedures of deliberation interfere with the mechanisms of group polarization. Based on a survey, people with either permissive or restrictive attitudes toward immigration were first identified and then invited to the experiment. The participants were randomly assigned to like-minded and mixed small-n groups. Each like-minded group consisted of only permissive or restrictive participants, whereas each mixed group consisted of four permissive and four restrictive participants. The like-minded treatment represents enclave deliberation, and the mixed treatment a ‘standard’ deliberative mini-public design. The main finding of our experiment is that people with anti-immigrant attitudes become more tolerant even when they deliberate in like-minded groups. Moreover, similar learning curves are observed in both treatments. Based on the results, we conclude that deliberative norms can alleviate the negative consequences of discussion in like-minded groups.

Keywords Enclave deliberation · Deliberative democracy · Group polarization · Group think · Experiment · Immigration attitudes

K. Grönlund (✉)

Social Science Research Institute, Åbo Akademi University, 20500 Åbo, Finland
e-mail: kimmo.gronlund@abo.fi

K. Herne

School of Management, Politics, University of Tampere, 33014 Tampere, Finland
e-mail: kaisa.herne@uta.fi

M. Setälä

Department of Political Science, University of Turku, 20014 Turku, Finland
e-mail: maija.setala@utu.fi

Introduction

Theories of deliberative democracy provide normative criteria for the evaluation of political discussion (Delli Carpini et al. 2004; Dryzek 2000). Moreover, the quality of public decisions can be expected to depend on the quality of democratic deliberation preceding decision-making. Deliberation can be defined as communication based on the merits of arguments, such as the sophistication of justifications and the generalizability of moral principles (e.g. Steenbergen et al. 2003). Recently, broader definitions of deliberation have been put forward in the literature. Along with rational argumentation, these definitions include other forms of communication such as rhetoric and narratives (see e.g. Mansbridge et al. 2010). Despite these developments, the idea of public reasoning among free and equal individuals remains in the core of the concept of democratic deliberation. Moreover, deliberative democrats agree that deliberation involves both intersubjective processes of exchanging arguments and internal processes of reflection based on these arguments (Goodin 2000).

One of the key features of deliberation is the *inclusion* of different viewpoints in the process of exchanging arguments. Indeed, the presence of conflicting viewpoints is often regarded as a necessary condition for deliberation. For example, Thompson (2008, 502) argues as follows: “If the participants are mostly like-minded or hold the same views before they enter into the discussion, they are not situated in the circumstances of deliberation”. However, the term ‘enclave deliberation’ has been increasingly used to refer to discussion among like-minded people. Cass Sunstein (2002, 2007, 2009) has emphasized the problems and risks related to enclave deliberation, most importantly group polarization. Other scholars (e.g. Karpowitz et al. 2009; see also Sunstein 2007, 76–77) have pointed out the importance of enclave deliberation for political articulation and mobilization, especially for those in disadvantaged positions. Indeed, Mansbridge (1994, 63), who was probably the first to use the term enclave deliberation, called for “enclaves of protected discourse and action” as an element of a just society.

In this paper, we study how the outcomes of deliberation vary depending on whether the deliberative discussion takes place among like-minded people or among people who disagree on a political issue. Our analysis is based on an experiment where citizens, drawn from a random sample, deliberated on immigration. The experiment was held in Finland in the spring of 2012. Although group composition is subject to manipulation in the experiment, the other ‘standard’ procedures of deliberative mini-publics are applied. These procedural features include balanced information provided to all participants as well as the use of moderators and discussion rules designed to encourage a process where people are exposed to different arguments and reflect on their own position in relation to them. We aim to look at the effects of group composition while holding the deliberative context constant. Our study therefore follows the recommendation made by Mutz (2006, 61) as we aim to disentangle the effects of group composition from the other aspects of the deliberative ‘package’.

The paper is organized in the following manner. First, we discuss the idea and mechanisms of enclave deliberation together with a literature review. After that,

four hypotheses are formulated. Third, the experimental procedure is described. Fourth, the data gathered in the surveys during the experiment are analyzed. Finally, conclusions are provided together with discussion on the findings.

Earlier research on enclave deliberation

Cass Sunstein (2002, 2007) has raised the question of the future of democracy if people only ever listen and speak to the like-minded. He addresses the problem of ‘group thinking’ which may arise when like-minded people discuss among themselves. It may lead to group polarization and an amplification of cognitive errors. Group polarization occurs when deliberation in a group of like-minded participants reinforces the attitudes and opinions prevailing in the group at the outset. Sunstein (2009, 3) defines polarization as follows: “[...] members of a deliberating group usually end up at a more extreme position in the same general direction as their inclinations before deliberation began”. Group thinking can also affect people’s factual beliefs. According to Sunstein (2007, 80–95, 140–143), enclave deliberation may lead to an amplification of cognitive errors, which means that biased or erroneous epistemic beliefs are corroborated. He also points out that large-scale misconceptions, or ‘informational cascades’ may come up in enclave deliberation because people just follow the cues provided by others in the absence of contrary evidence.

There are different mechanisms contributing to polarization when opinions in a group are biased at the outset. More precisely, two types of mechanism have been identified; namely social comparison and persuasive arguments (Farrar et al. 2009, 616; Isenberg 1986; Sunstein 2002, 179–180). Social comparison refers to the tendency of individuals to act in order to win social acceptance from other members of the group. In order to be accepted, individuals need to process information of how other people present themselves, and adjust their own behavior accordingly (Isenberg 1986, 1142). Individuals may act in different ways in order to be perceived favorably by other group members. First, they may try to adjust their opinions according to the view which seems to dominate the group.¹ Social psychological experiments have also demonstrated that group pressures work in the way that people tend to conform to the views of the majority (Asch 1948). Second, social comparison may also make people emphasize their difference from others to the valued direction (Isenberg 1986, 1142). In other words, individuals may take positions which are more extreme in comparison to the views dominating the group at the outset.

The other mechanism behind group polarization, persuasive arguments, is simply based on the idea that individuals are convinced by the contents of arguments put forward in the group. Consequently, if arguments heard in a group are biased in one direction, there is likely to be a further shift to this direction. Group polarization is likely to be reinforced by biases in information processing. ‘Confirmation bias’ is a

¹ The so-called ‘spiral of silence’ means that, in fear of social isolation, people who think that their political views are unpopular remain silent in the group (Noelle-Neumann 1993, 201–202).

well-established phenomenon which means that people are inclined to seek information confirming their prior beliefs and to disregard information against them (Mercier and Landemore 2012, 251). More generally, motivated reasoning refers to a variety of cognitive and affective mechanisms which lead individuals to arrive at the conclusions they want to arrive at (Kunda 1990). In a group of like-minded people, individual biases in information processing and reasoning are not checked by arguments put forward by individuals supporting conflicting views. Opinions are likely to polarize because individuals only hear arguments supporting their own position—in fact; they may even hear new arguments in support of it.

The negative consequences of discussion among like-minded people have been confirmed by social psychological studies. Sunstein himself (2007, 60–62) provides some experimental evidence on group polarization. He puts forward a summary of social psychological studies showing that groups tend to move toward the direction of the position initially dominating the group (Sunstein 2009, 161–168). The discussion topics of the described experiments range from jury decisions to risk taking and militarism and pacifism. In a recent study on political discussion, Jones (2013) found evidence on the polarization of opinions, especially among the Republicans, in a partisan workplace environment.² Some studies on group polarization have not even involved proper discussion. For example, Lee's (2007) study showing that group polarization is correlated to group identification is based on computer-mediated communication and not actual discussion between the group members.

Sunstein's account on group polarization emphasizes the original dispositions of those who discuss and the biases in the argument pool. However, the studies he discusses do not represent experiments on democratic deliberation understood as a specific form of discussion where certain standards of reasoning and argumentation are followed. So-called deliberative mini-publics (Goodin and Dryzek 2006) involve procedures enhancing democratic deliberation. Most importantly, participants of mini-publics receive balanced information on the issue and group discussions are moderated. Results from deliberative mini-publics provide a different picture from that of Sunstein's: Groups de-polarize rather than polarize, people learn during deliberation and their misperceptions are corrected (e.g. Luskin et al. 2002; Setälä et al. 2010; Grönlund et al. 2010). Moreover, people learn about facts supporting views which they initially disagreed with (Andersen and Hansen 2007). These results may be explained by the fact that the inclusion of different viewpoints is ensured in deliberative mini-publics which therefore do not represent deliberation in like-minded groups. This explanation suggests that group composition may be a crucial determinant of deliberative outcomes (Mendelberg and Karpowitz 2007).

Sunstein (2009, 48) argues as follows: “When groups contain equally opposed subgroups, do not hold rigidly to their positions, and listen to one another, members will shift toward the middle; they will depolarize. The effect of mixing will be to produce moderation”. Mercier and Landemore (2012, 253) argue that in genuine deliberation, the biases present in individual reasoning are checked by biases in arguments of individuals representing different viewpoints. Moreover, without a

² Mutz and Mondak (2006), on the other hand, use survey evidence to demonstrate that the workplace commonly provides an environment for cross-cutting political discourse.

clearly dominant view in the group, people are likely to accommodate their arguments in ways which could appeal to people representing conflicting viewpoints. Indeed, Mercier and Landemore argue that collective processes of deliberation are likely to be the most effective remedies for biases in information processing and reasoning.

Hypotheses

In this article, we analyze the impact of group composition on the outcomes of deliberation. More specifically, we compare the outcomes of deliberation in like-minded groups with those groups where the participants' opinions are divided. The analysis is based on an experiment where citizens were invited to deliberate on immigration policy. Based on earlier theoretical and empirical findings, we test four hypotheses which relate to opinion and knowledge changes. H_{1a} and H_{1b} concern opinions, whereas H_{2a} and H_{2b} address cognitive errors.

- H_{1a} Deliberation in like-minded groups leads to a polarization of opinions.
- H_{1b} Deliberation in mixed groups de-polarizes opinions.
- H_{2a} Deliberation in like-minded groups amplifies cognitive errors.
- H_{2b} Deliberation in mixed groups corrects cognitive errors.

We are interested in finding out whether the procedures enhancing deliberation in mini-publics, such as balanced information, discussion rules and moderation can restrain the negative consequences of enclave deliberation. *Therefore, it is possible that there will be less pronounced differences between the treatments than anticipated in the hypotheses.* Although we assume that the dynamics and outcomes of deliberation differ depending on the group composition, it should also be acknowledged that the set-up of deliberative mini-publics may foster deliberative forms of communication even in groups of like-minded people. For example, Farrar et al. (2009) found that the group composition has little impact on deliberative outcomes in their randomized experiments. Moreover, the authors point out several factors which can neutralize group composition effects, including the types of topics discussed and the presence of varying viewpoints.

In deliberative mini-publics, information provided to participants can be expected to broaden the set of arguments put forward in discussions. It may also create a common pool of arguments which may lessen polarization (see e.g. Isenberg 1986, 1148). Moreover, the use of moderators and the application of specific rules for discussion are likely to enhance a free and equal exchange of viewpoints as well as an evaluation of arguments based on their merits. As pointed out by Kunda (1990), the tendency of individuals to move toward motivated reasoning is constrained by their ability and willingness to construct reasonable justifications for their conclusions. The set-up of a deliberative mini-public, emphasizing processes of reasonable justification as 'the name of the game', can be expected to encourage reasoning and argumentation about the *pros* and *cons* concerning the issue. This type of reasoning can be enhanced further if participants

adopt argumentative strategies which help to compensate the biases in the argument pool, for example, by acting as ‘devil’s advocates’.

Experimental procedure

The topic of the deliberation experiment was immigration policy which is a contested and debated issue in Finland. Over the last years, the populist right wing Finns party,³ in particular, has kept the immigration issue on the agenda. The main purpose of our experiment was to compare deliberation in two types of small groups: (1) groups consisting of like-minded people, and (2) groups consisting of people having different opinions on immigration. Those who indicated willingness to take part in the experiment were randomly assigned to like-minded groups, mixed groups, and a control group. Subjects in the first two groups took part in the deliberation event, whereas the control group only filled in three mail-in surveys.

The participants’ opinions were measured before and after deliberation. Respondents with negative attitudes to immigration formed a *con* enclave, whereas respondents with a positive view on immigration formed a *pro* enclave. Within enclaves, subjects were randomly assigned into two treatments and a control group.

A short survey (T1) was first mailed out to a simple random sample of 12,000 adults in the region of Turku/Åbo. The sample was provided by the official population registry of Finland. Of the addressed sample, 39 % ($n = 4,681$) responded to the survey. T1 consisted of 14 items measuring the respondents’ attitudes on immigration. The questions were first pilot tested with students at two universities in order to measure the appropriateness of the questions for the purpose of the experiment. All survey items worked well both in the pilots and in the actual survey conducted among the random sample (T1). In the surveyed sample, all 14 items loaded on one single factor and Cronbach’s Alpha of the sum variable reached 0.94. Therefore, we concluded that the questions measured attitudes toward immigration on a one-dimensional scale and constructed a sum variable of the 14 items. Each item was first recoded into a scale from 0 to 1, so that 1 indicates the most immigration-friendly attitude. Thus, the index can vary between 0 and 14. The questions are listed in Appendix 1. Figure 1 shows the initial dispersion of attitudes among those respondents ($n = 3,232$) who allowed further contact from the research group.

The histogram shows that the initial opinions almost followed the normal distribution. Thus, we felt confident to use the sum variable as a ground for creating the *con* and *pro* enclaves. Since the design of the experiment required people with clear views on the immigration issue, we excluded moderates, i.e. respondents whose opinions on immigration were close to the median value of the frequency distribution ($n = 631$).⁴

³ The Finnish name of the party is *Perussuomalaiset* and it is sometimes referred to as the ‘True Finns party’. The party won a landslide victory in the Parliamentary election of 2011 with 19.1 % of the votes.

⁴ Those whose value for the sum variable was >8.3 were included in the *pro* enclave, and those whose value was <6.7 were included in the *con* enclave. The cutoff points were chosen based on the distribution of the opinions in the sample. The median value is 7.61 and the arithmetic mean 7.28 with a standard deviation of 2.98. The cutoff points grouped deciles 1–4 into the *con* enclave and deciles 7–10 into the *pro* enclave. Thus, the fifth and sixth deciles were left out from further contacts.

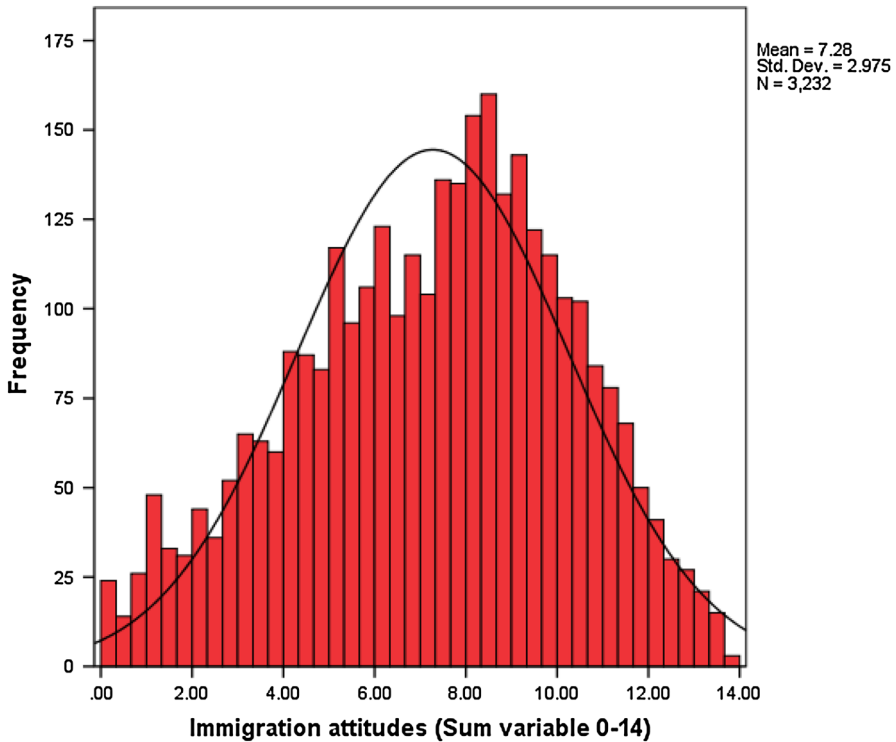


Fig. 1 The dispersion of initial attitudes on immigration among the respondents

The second survey (T2) with 37 items and an invitation to take part in a deliberation (and a separate debriefing) event was sent to 2,601 people who qualified as members of either the *con* or the *pro* enclave. At this point, it was clearly stated that the deliberation event was an integrated part of the research project and that a response to the survey meant a preliminary agreement to take part in it. Furthermore, it was clarified that only a part of those who volunteered could be included in the deliberation event and that the choice would be made by lot. Each participant who completed all stages of the project was compensated. A gift certificate worth 90 Euros was given to each participant of the deliberation and debriefing events and 15 Euros to those whose task was only to fill in the surveys (i.e. the control group).

Eventually, 805 people volunteered, and 366 were invited to take part in the deliberation event. The target sample was 256 participants, which would have allowed for 32 small groups of eight participants (eight *pro* like-minded, eight *con* like-minded and 16 mixed groups). Stratified sampling was used in order to guarantee representation in terms of the *pro* and *con* enclaves as well as age and gender. Random sampling was used within the two strata. Unfortunately, the target of 256 deliberators was not achieved and 207 people showed up.⁵ Especially people

⁵ A more thorough analysis of the attrition process can be found in Karjalainen and Rapeli 2015.

in the *con* enclave tended to abstain at this final stage, even though there were no indications of this kind of a bias at the earlier stages of the recruitment process. Figure 2 shows the phases of the recruitment process.

Since some of the invited participants with anti-immigrant attitudes dropped out at the final recruitment stage, we wanted to check if the sample of people turning out to deliberate was skewed when it comes to attitudes. In Table 1, comparisons between the preliminary invited sample ($n = 2,601$), the initially volunteered respondents ($n = 805$), the invited ($n = 366$) and the actual participants ($n = 207$) are made within the two enclaves. It can be seen that the participants in the *con* enclave were slightly more moderate, i.e. less anti-immigrant, compared to the whole enclave at earlier stages. In fact, the difference in opinions in the *con* enclave between the participants ($n = 86$) and the ones who did not show up ($n = 97$) is statistically significant at the 0.01-level. In other words, it was harder to attract people with the most anti-immigrant opinions to present their views in a deliberative event. In the *pro* enclave, the participants were slightly more liberal than the mean of the whole enclave at earlier stages. This difference is not, however, statistically significant.

For the deliberation event, random allocation was used within the *con* and *pro* enclaves. There were 10 *pro* like-minded groups, five *con* like-minded groups and 11 mixed groups.⁶ Table 2 displays the assignment into the four types of groups and shows the number of individuals within each cell. The control group consisted of 369 people who were initially willing to take part and who returned each of the surveys T1, T2 and T4. The socio-demographics of the participants and the control group are shown in Appendix 3.

The phases of the experiment are described in Table 3. The deliberation event took place during one weekend, 31 March–1 April 2012. Each participant took part only during 1 day, either on Saturday or Sunday. Each day, the event followed the same procedures and lasted from 9.30 a.m. until 3 p.m. The day started with a short 15-item quiz (T3) measuring knowledge related to immigration and general politics. The quiz was composed of four immigration-related items that were included in the briefing material, six issues that were not included in the briefing material, and five issues related to general political knowledge. After the knowledge quiz, the participants were briefed about some basic facts related to immigration in Finland. The briefing was designed to be balanced, focused on basic facts, and it was presented as a slide show in an auditorium to all participants. It consisted of statistical data on migration from and to Finland over the years as well as the number of immigrants by country of origin in Finland. Furthermore, it explained the official processes required for legal immigration and statistics about different grounds for acquiring a residence permit in Finland. The material ended with a short immigration-related glossary. All politically controversial topics, e.g. statistics about crime and unemployment, were left out. A copy of the information material was also handed out to each participant.

⁶ One of the mixed groups was a Swedish-speaking group allowing participants to talk in their mother tongue.

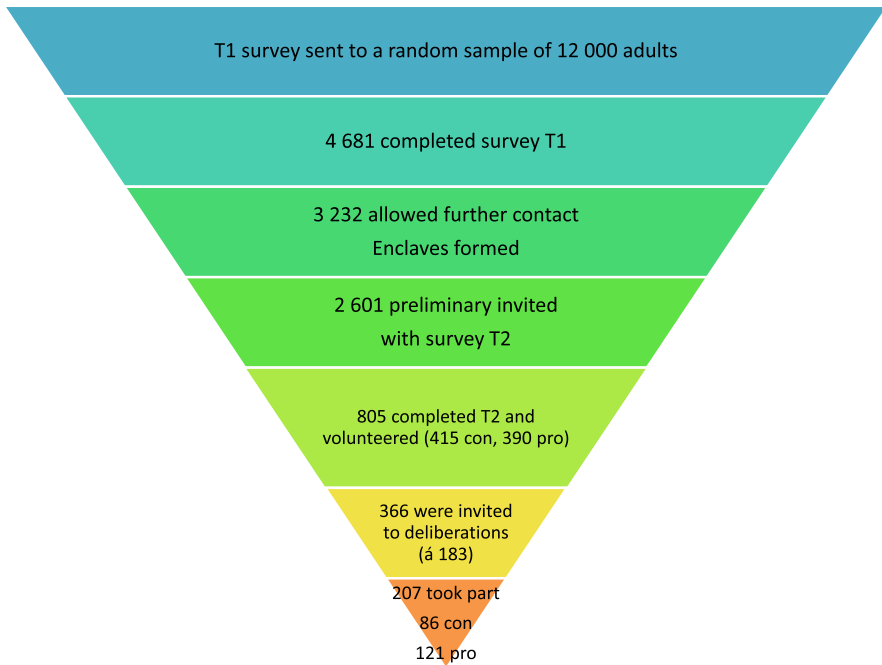


Fig. 2 Recruitment and attrition

Table 1 A comparison between the preliminary invited, the volunteered, the invited and the participated individuals

	N			T1 opinions on immigration	
	All	Con	Pro	Con	Pro
Preliminary invited	2,601	1,304	1,297	4.28	10.15
Initially volunteered	805	415	390	4.25	10.17
Invited	366	183	183	4.33	10.24
Participated	207	86	121	4.68	10.26

Table 2 Group formation

		Randomization	
		Like-minded	Mixed
Stratification	Con	Con like-minded (n = 42)	Con mixed (n = 44)
	Pro	Pro like-minded (n = 77)	Pro mixed (n = 44)

Table 3 The phases of the experiment

Pre surveys (January 2012)
1. Short survey to form enclaves (T1)
2. Second survey with invitation (T2)
The deliberation event (March 31–April 1, 2012)
3. Knowledge quiz (T3)
4. General instructions and briefing on the immigration issue
5. Deliberation in small-n groups (4 h, incl. 45 min lunch break)
6. Survey measuring opinion and knowledge changes and experiences of the event (T4)
Debriefing (April 20, 2012)
7. A follow up survey measuring the stability of opinion changes (T5)
8. Debriefing of the study for participants

After the briefing, facilitated small-n group deliberations began. Each mixed group consisted of exactly eight participants, of which four were randomly selected from the *con* enclave and four from the *pro* enclave. While it was seen as crucial to achieve an exact balance between the two enclaves in each mixed group, small variation in group size was allowed in the like-minded treatment. This was due to attrition at the last stage.⁷

The group discussions lasted for 4 h, including a lunch break of 45 min. The group discussions ended with a survey (T4) repeating the questions in T1, T2 and T3, apart from socio-economic background variables. The survey also included questions on the participants' experiences of the deliberation event.

In each group, a trained moderator facilitated the discussion. A written description of the rules of the discussion was handed out to the participants, emphasizing respect for other people's opinions, the importance of justifying one's opinions and openness to others' points of view. The moderators also read aloud these rules.

In the beginning of the group discussion, each group member put forward a theme related to the immigration issue which they wished to be discussed. The moderator wrote these themes down on a blackboard. The proposed themes covered issues such as employment-based immigration, humanitarian-based immigration, acculturation, multiculturalism, unemployment, crime and security, language and education, immigration attitudes, and the cost of immigration. There were no major differences between the themes put forward by the participants in the *pro* and *con* enclave. However, none of the *con* participants suggested the theme of immigration attitudes, i.e. prejudices and racism, as a discussion topic. After the round of introducing discussion themes, free discussion on the proposed themes followed. The moderators interfered only if any of the group members dominated or completely withdrew from the discussion. Furthermore, the moderator could put

⁷ Therefore, of the 15 like-minded groups, nine had eight participants, whereas three groups consisted of nine, two groups of seven and one group of six participants.

forward a theme for discussion from those written down on the board in case the discussion paused. Discussions in both like-minded and mixed groups lasted throughout the whole period of time. However, there were slight differences in discussion activity as there were, on average, more speech acts (70.8) in the mixed treatment than in the like-minded treatment (50.7).

Results – hypotheses testing

In this section we present the main results of the experiment, i.e. the development of opinions and knowledge in like-minded and mixed groups. The statistical significance of potential differences and changes is determined through *t*-tests. We compare both the development of opinions and knowledge according to the four groups achieved by the combination of enclave and treatment (see Table 2). The comparisons are mainly done within- subjects (paired pre- and post-test), but also between-subjects testing when applicable.

First, we test the hypotheses on the effect of treatment on opinions. H_{1a} states that a polarization of opinions occurs in like-minded groups, whereas H_{1b} assumes that the opposite occurs in mixed groups. Table 4 demonstrates the development of opinions in the course of the experiment. The comparisons are made within enclaves and treatments, as well as in the control group. We compare opinions before (T1) the event, after deliberation (T4) and in the follow-up survey (T5) 3 weeks after the event. In the control group, the measurements were done before (T1) and after (T4) deliberation.

There were three statistically significant opinion changes among those participants who took part in the deliberation event. All of these are in the direction of a more liberal attitude toward immigration. The most prominent change occurred among the participants of the *con* enclave deliberating in mixed groups. Here, the initial mean on the sum variable was 4.33, which increased to 6.12 after the

Table 4 Opinion changes by enclave and treatment as well as in the control group

	Before (T1)	After (T4)	Change	<i>p</i>	Follow up (T5)	Change (T5 – T4)	<i>p</i>	N
Con like-minded	5.05	5.73	0.67	**	6.15	0.42	*	42
Con mixed	4.33	6.12	1.80	***	6.32	0.20		44
Con control group	4.21	4.70	0.49	***	–	–		194
Pro like-minded	10.44	10.73	0.29	*	10.60	–0.12		77
Pro mixed	9.95	9.90	–0.06		10.04	0.14		44
Pro control group	10.06	9.55	–0.51	***	–	–		175

Mean values for the sum variable measuring opinions on immigration

The variable can vary between 0 (against) and 14 (in favor)

* *p* < 0.05; ** *p* < 0.01; *** *p* < 0.001

experiment. The increase of 1.8 units is significant at the 0.001 level and corroborates H_{1b} . On average, the participants of the *pro* enclave deliberating in mixed groups did not, contrary to H_{1b} , move toward the middle when deliberating with “the other side”. Thus, depolarization in the mixed treatment was unilateral, only persons with anti-immigrant attitudes shifted toward the mean. Those initially permissive toward immigration did not alter their opinions.

Moving on to the like-minded treatment where polarization is most likely to occur, Table 4 shows that the *con* like-minded groups did not polarize in comparison with their initial opinions. On the contrary, participants in the *con* like-minded groups became more permissive toward immigration as a result of deliberation. The change of 0.67 units is not as large as among the *con* participants in the mixed groups, but still significant at the 0.01 level. This development works against H_{1a} . In the *pro* enclave, the like-minded groups show a barely statistically significant (0.05) mean change of opinions. These groups polarized slightly, according to the assumption in H_{1a} .⁸ Furthermore, the overall patterns found at the aggregate level were confirmed in a separate group-by-group analysis. None of the small groups behaved in a deviant manner. In 9 (out of 11) of the mixed groups, in 4 (out of 5) of the *con* like-minded groups, and in 7 (out of 10) of the *pro* like-minded groups, the change in opinions toward immigration was positive from T1 to T4. Looking at the follow-up survey T5, we can see that the only statistically significant change between deliberation and the follow-up survey was a continued tendency among the *con* like-minded groups to become more tolerant toward immigration. At T5, the participants of the *con* like-minded treatment had shifted from 5.05 before deliberation to 6.15. This change corresponds to a 1.1-unit increase on the 14-item scale ($p < 0.001$).

Moving on to comparing the treatment groups with the control group, it can be seen that attitudes toward immigration changed also in the control group. Within the control group, the *con* enclave became slightly more permissive (change 0.49), whereas the *pro* enclave became slightly more critical (change -0.51). In other words, participation in a three wave (T1, T2, and T4) panel study on immigration seems to have led to a de-polarization of opinions among the control group. It may be the case that people who responded to the survey became more aware of the immigration issue even though they did not participate in the deliberation event. They may, for example, have sought more information on their own and reflected on it. Another possible explanation is the statistical phenomenon known as ‘regression to the mean’. It occurs when the same sample is measured twice, especially when the used measurement is not accurate. In survey research, measurement errors are bound to occur and observations with the most extreme values tend to regress towards the mean at the second measurement (Torgerson and Torgerson 2008, 10–15). When designing the experiment, we tried to minimize measurement errors

⁸ The non-parametric Wilcoxon signed-rank test confirms the obtained results, T1–T4, in the four group types (for *con* participants in like-minded groups $p = 0.005$; for *con* in mixed groups $p = 0.000$; for *pro* in like-minded groups $p = 0.032$; and for *pro* in mixed groups $p = 0.899$). Likewise, the Mann–Whitney test confirms the overall pattern about the differences between treatments at T1 (for the difference between *con* participants in like-minded and mixed groups $p = 0.021$; between *pro* participants in like-minded and mixed groups $p = 0.087$), and at T4 (between *con* participants in like-minded and mixed groups $p = 0.294$; between *pro* participants in like-minded and mixed groups $p = 0.002$).

by creating an index that consists of 14 items. Thus, it should be a more accurate measurement of opinions toward immigration than a single variable or an index consisting of only a few items. Still, survey questions with Likert scales do not provide an exact measure and the possibility of regression to the mean both in the control group and the treatment groups cannot be ruled out. However, the fact that the opinion changes in the control group did not follow the same patterns as the opinion changes in the experimental groups suggests that deliberation had a genuine effect on opinions. It is also important to note that both the experimental and the control group was formed randomly from those willing to participate in the deliberation event, indicating that a self-selection bias cannot account for the differences observed between the two groups.

Table 4 also shows a comparison within enclaves at T1, which helps to trace possible initial differences between the subjects who were randomly allocated into the two treatments within both enclaves. Despite random assignment, there were some differences between the participants in the like-minded and the mixed treatment in both enclaves. In the *con* enclave the participants of the like-minded treatment were somewhat more moderate than the participants in the mixed treatment. In the *pro* enclave, the opposite was the case, i.e. the like-minded treatment consisted of more liberal participants than the mixed treatment. It is hard to decipher whether this initial division had any influence on the outcome within the treatments. However, in order to understand how extremes and moderates behaved within both treatments, we have conducted additional analyses further below (Table 6).

Next, in order to test H_{2a} and H_{2b} we analyze knowledge change in the course of the experiment. H_{2a} suggests that deliberation in like-minded groups amplifies cognitive errors, whereas H_{2b} anticipates that deliberation in mixed groups corrects cognitive errors. The knowledge questions were grouped in three subsets. First, there were six questions pertaining to immigration where information was given in the beginning of the deliberation event. Second, there were four questions on immigration where information was not given by the organizers. Third, there were five questions measuring general political knowledge. In Fig. 3, we look at the learning effects by treatment and enclave. The figure only includes the ten items related to immigration knowledge (for a detailed development of all knowledge items, see Appendix 2).

The participants learned a lot during the experiment. The obtained learning effects were large and similar in all four types of groups and we can conclude that neither treatment nor initial attitudes toward immigration had an effect on the learning curve. We can also see that the pre-deliberation knowledge shares were quite similar across enclaves and treatments. For all participants, the mean share of correct answers increased from 43 to 63 %, and the information gains were recorded for those questions where information was given at the event (see Appendix 2). This indicates that knowledge gains occurred both in mixed and like-minded groups to a similar degree, working against H_{2a} . Initially, there were small differences within enclaves between the subjects who were randomly assigned to the like-minded versus mixed treatments. These differences were not statistically significant, and neither were the differences within enclaves at T4.

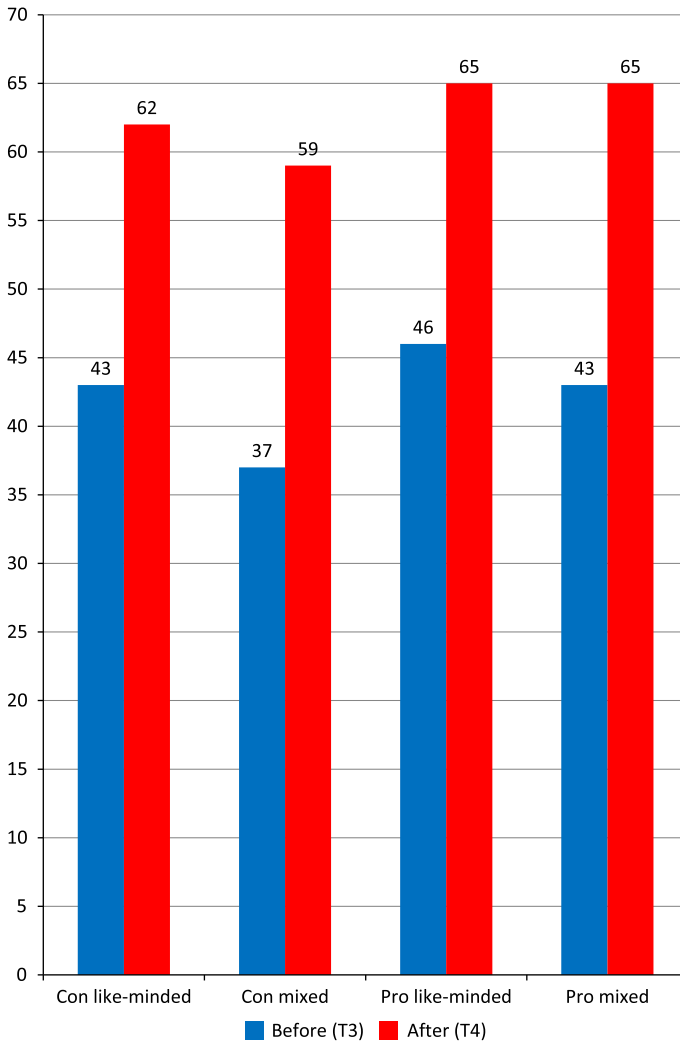


Fig. 3 Knowledge gains according to enclave and treatment. Shares of correct answers

Within the set of questions relating to immigration where no information was provided by the organizers there were two open-ended questions (questions 9 and 10 in Appendix 2),⁹ which can be used to examine the hypotheses concerning cognitive errors. These questions pertained to the level of unemployment among immigrants (correct answer 27 %) and the level of social security benefits received by an unemployed immigrant (correct answer 757 Euros per month). It can be assumed that negative attitudes toward immigration are, especially, related to people's perceptions of social problems and costs caused by immigration. Therefore, it may

⁹ All other knowledge items were put forward as multiple choice questions with four alternatives.

be assumed that those who have negative attitudes toward immigration might overestimate both the level of unemployment and social security benefits, and the opposite could be the case among supporters of immigration.

Whereas the coding of the open-ended questions in Appendix 3 follows the binary logic of ‘correct’ and ‘non-correct’¹⁰ answers, we have also examined the distance of each respondent’s answer from the correct answer. When looking at the responses to the open questions before deliberation (T3), it turns out that the *pro* enclave participants, in particular, tended to underestimate the unemployment rate and the level of social security (42.2 % of the participants in the *con* enclave and 53.3 % in the *pro* enclave underestimated the unemployment rate. The level of social security benefits for immigrants was underestimated by 67.5 % in the *con* enclave and 68.1 % in the *pro* enclave). There were clearer differences between enclaves when we look at overestimation, which gives some support to the assumption that attitudes toward immigration are related to the perceptions of the costs of immigration. Namely, 33.7 % of the participants in the *con* enclave overestimated the unemployment rate as opposed to 23.3 % in the *pro* enclave. The level of social benefits was overestimated by 21.7 % of the participants in the *con* enclave as opposed to 10.1 % in the *pro* enclave. However, the responses to the open-ended questions after deliberation (T4) do not support the hypotheses on the amplification of cognitive errors. The differences in the under- or overestimation of the unemployment rate and social security benefits are not statistically significant when comparing the enclaves and treatments with each other.

To sum up, none of our hypotheses has gained clear support so far. We do not trace polarization effects, with the partial exception of the *pro* participants in like-minded groups, and the differences between the like-minded and the mixed treatments are modest so far. Moreover, the obtained learning curves are similar in all four types of groups, and we cannot see any results supporting the assumption of increased cognitive errors in the like-minded treatment. The major conclusion is that as a result of deliberation most subjects became more tolerant toward immigration, and that at the group level no one became less tolerant. In the next part of the empirical analysis, we will try to disentangle the observed group level results by looking at individuals.

Further analyses on the impact of deliberation

In order to understand the scope of opinion changes at the individual level, Table 5 focuses on participants who changed sides as a result of deliberation. By changing sides we mean a shift from being initially against immigration to becoming pro-immigration as a result of deliberation, or vice versa. The threshold for changing sides is set at 7.5 on the sum variable, i.e. in the middle of the original cutoff points for forming the *con* and *pro* enclaves. If a person in the *con* enclave moved above

¹⁰ The intervals of acceptance for correct answers were defined as follow: 24–30 % for the unemployment rate, 700–800 EUR for the integration assistance. The intervals were chosen by taking into account the dispersion of answers and the nature of the question. These open questions proved to be difficult but we did not want to stretch the category of ‘correct answer’ too far from the correct numbers.

7.5 after deliberation, (s)he is considered to have changed sides; a person in the *pro* enclave should have moved below 7.5 in order to have changed sides concerning opinions on immigration. Table 5 compares the participants' pre and post deliberation attitudes. Moreover, the post deliberation attitudes are analyzed both at the end of the deliberation event (T4) and in the follow-up survey (T5).

Changing sides occurred almost exclusively among persons who were initially against immigration. This corroborates the earlier finding that the largest opinion changes took place among the participants with anti-immigration attitudes. At the end of the deliberation event (T4), 20 persons belonging to the *con* enclave had become permissive toward immigration; the sum variable for immigration attitudes had, for their part, exceeded the value of 7.5 on the 0–14 point scale. A majority of these, 14 people, were subject to the mixed treatment. When people with anti-immigrant attitudes faced counter-arguments, many of them became clearly more positive toward immigration. However, also six people in the *con* like-minded groups changed sides at T4. No one in the *pro* enclave became restrictive toward immigration as a result of deliberation.

Moving on to the follow-up survey, the trend to more permissive attitudes continues among the participants of the *con* like-minded groups. In the period between the experiment and the follow-up survey, five additional people in *con* like-minded groups had become supporters of more immigration. To sum up, 14 out of 44 *con* participants in mixed groups changed sides as a result of deliberation (13 at T5 as one person had shifted back), and 11 out of the 42 *con* participants deliberating in like-minded groups had changed sides at the follow-up survey. Altogether 24 of the total of 86 participants in the *con* enclave changed sides as a result of the experiment. Considering that the threshold for becoming tolerant was drawn at 7.5, and not at 7 which would be at the middle of the scale, we can conclude that the results show a clear shift toward a more liberal position at the individual level. *More than every fourth participant with anti-immigration attitudes became permissive toward immigration as a result of deliberation.*

Did the fact that some participants were more moderate and some more extreme have an effect on the obtained results? We have earlier acknowledged that even though random assignment was used, it produced a somewhat biased division between the two treatments. Participants in the like-minded treatment were, by chance, somewhat closer to the mean position on the 14-item scale than the participants in the mixed treatment. This was true for both enclaves, which raises the question of whether the observed opinion changes reflect the fact that moderates were more inclined to change toward the mean than participants with more extreme views. In order to test this, we divided the sample into four new groups. Participants with the label '*con* extreme' scored less than 5 on the scale at T1, whereas participants labeled '*con* moderate' had values between 5 and 6.7 (the cutoff point). Among the liberal participants, the *pro* moderates scored between 8.3 and 9.99, whereas the '*pro* extremes' scored at least 10. Table 6 compares these four groups of subjects within the two treatments.

The dependent variables in Table 6 consist of the amount of opinion change and learning on immigration issues in the course of the experiment, as well as a variable measuring the discussion activity of the participants. Discussion activity was coded as a relative measure showing how much each individual talked within their group.

Table 5 Did the participant change sides during deliberation (compared to T1)?

	Con		Pro		Total
	Like-minded	Mixed	Like-minded	Mixed	
At T4					
Yes	6	14	0	0	20
No	36	30	77	44	187
Total	42	44	77	44	207
At T5					
Yes	11	13	1	0	25
No	31	31	76	44	182
Total	42	44	77	44	207

The variable was coded from the transcriptions of the audio recorded deliberations in each group and measures the number of characters spoken by a person in relation to the total number of characters spoken in the group, excluding the moderator.¹¹ Table 6 shows that the observed changes in the *con* enclave were not caused by moderates becoming more tolerant. On the contrary, the participants with more extreme anti-immigrant opinions shifted clearly more in a liberal direction than the moderates did. In the mixed treatment, the shift among extreme *con* participants was over 2 units on the 14-item scale, and also in the like-minded treatment, it was over 1.2 units. Especially in the like-minded treatment, participants with extreme anti-immigrant opinions were the ones who drove the group mean toward a more liberal direction.

Within the *pro* enclave, the observed minor polarization was caused by the moderates becoming more tolerant; they shifted 0.67 units upwards, whereas the participants with extremely tolerant views stayed put. This difference is statistically significant at the 0.01-level. Moving on to knowledge change and discussion activity, Table 6 shows that there were no differences between the moderate and extreme participants in either enclave or treatment. The observed knowledge gains were similar in each group and the relative activity of the participants did not differ within treatments and enclaves according to how extreme their baseline views were in relation to each other.

In order to gauge the impact of knowledge gains on opinion change, we also divided the participants into those who learned and those who did not learn about immigration in the course of the deliberation event.¹² This analysis is relevant in

¹¹ The number of characters was chosen instead of the number of words since the Finnish language consists of long words and no prepositions are used. Instead, the language uses cases which adds characters to words. In practice, the relative numbers of words and characters were highly collinear ($r_{xy} = 0.84, p < 0.01$).

¹² Most participants learned correspondingly to one item of the ten items measuring knowledge related to immigration, which makes dichotomization somewhat tricky. Therefore, we label participants who scored two or more correct answers after deliberation as learners ($n = 85$), whereas the rest are coded as non-learners ($n = 122$). Because of the ceiling effect, also the two most knowledgeable persons who scored eight correct items at T3 were coded as learners (one of them learned one more item during deliberation, whereas the other remained at 8).

Table 6 Comparisons between participants with moderate and extreme views within enclaves and treatments

Type of participant →	Opinion change		<i>p</i> (diff.)	Knowledge change		<i>p</i> (diff.)	Discussion activity %		<i>p</i> (diff.)
	Extreme	Moderate		Extreme	Moderate		Extreme	Moderate	
Con like-minded	1.24	0.36	α	0.40	0.37		13.94	11.30	
Con mixed	2.09	1.34	α	0.37	0.53		10.74	11.59	
Pro like-minded	-0.01	0.67	**	0.35	0.41		13.61	12.99	
Pro mixed	-0.28	0.10		0.44	0.44		15.14	12.45	

Variables: Opinion change between T1 and T4, Change in immigration knowledge between T3 and T4 (coded into a scale 0–1), relative discussion activity within the group

Extreme con like-minded (*n* = 15), Moderate con like-minded (*n* = 27), Extreme con mixed (*n* = 27) Moderate con mixed (*n* = 17), Extreme pro like-minded (*n* = 43),

Moderate pro like-minded (*n* = 34), Extreme pro mixed (*n* = 18), Moderate pro mixed (*n* = 26)

α *p* < 0.1; * *p* < 0.05; ** *p* < 0.01; *** *p* < 0.001

order to rule out the possibility that the perceived opinion changes were caused by the briefing given by the organizers. Each participant is placed in their enclave and treatment. The main variable of interest is opinion change, but as in Table 6, we also look at the discussion activity of each subject. Table 7 displays the impact of learning.

Table 7 shows that learning was not a key factor driving opinion change. Within the *con* enclave, the difference in opinion change between learners and non-learners is not significant in either treatment. There is only one statistically significant difference in the table. In the *pro* enclave, those participants who were subject to the like-minded treatment *and* did not learn became more extreme. This seems to give some support to the existence of a link between cognitive errors and polarization, as suggested by Sunstein. Further, those participants in the *pro* enclave who learned in the course of deliberation did not polarize. Regarding discussion activity, we can see that there were no differences in activity between those who learned and did not learn within any of the four group types. This indicates that the deliberations in the small-n groups were balanced and equal. To sum up, the additional analyses conducted in Tables 5, 6, and 7 seem to support the initial interpretations we made in conjunction with hypotheses testing.

Discussion

The results from the experiment do not show systematic patterns of group polarization in the like-minded groups. H_{1a} and H_{1b} concerning attitude changes gain only partial support. Those people in the *con* enclave who deliberated in like-minded groups did not become more extreme: on the contrary, they became more permissive toward immigration.¹³ In the mixed groups, participants in the *con* enclave became more permissive, whereas participants in the *pro* enclave did not become more critical toward immigration. Depolarization therefore occurred in the mixed treatment but it concerned the *con* enclave only. It is also notable that we did not find any clear indications of an amplification of cognitive errors in the like-minded treatment. H_{2a} and H_{2b} anticipated differences in learning between the treatments. Contrary to our hypotheses, participants assigned to mixed groups did not learn more than participants assigned to like-minded groups. Most participants learned to a substantial degree but this was mostly a result of the information given to them at the beginning of the deliberation event. Still, an interesting fact is that opinion polarization only occurred among those *pro* enclave participants in the like-minded treatment who did not learn in the course of deliberation. This group became even more tolerant (Table 7), which draws attention to the connection between non-learning and opinion polarization (c.f. Sunstein 2007, 80–95, 140–143).

One possible explanation to our findings on attitude change could be social desirability. Despite the increase of anti-immigration rhetoric in the Finnish political discourse, expressions of anti-immigration attitudes are still likely to face public

¹³ In this respect our results are similar to those received by Sanders (2012).

Table 7 Comparisons between participants who learned and did not learn within enclaves and treatments

Type of participant →	Opinion change		<i>p</i> (diff.)	Discussion activity %		<i>p</i> (diff.)
	Learned	Did not learn		Learned	Did not learn	
Con like-minded	0.57	0.75		14.47	10.73	
Con mixed	1.51	2.01		11.09	11.05	
Pro like-minded	−0.01	0.47	*	13.72	13.10	
Pro mixed	−0.28	0.13		13.56	13.52	

^a $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Variables: Opinion change between T1 and T4, relative discussion activity within the group

Learned con like-minded ($n = 17$), Did not learn con like-minded ($n = 25$), Learned con mixed ($n = 19$), Did not learn con mixed ($n = 25$), Learned pro like-minded ($n = 29$), Did not learn pro like-minded ($n = 48$), Learned pro mixed ($n = 20$), Did not learn pro mixed ($n = 24$)

disapproval. It is possible that participants felt pressure to provide politically correct liberal answers, i.e. that social desirability influenced their survey responses. Social desirability could influence participants' answers both when they respond to surveys at home and during the deliberation event itself.

There are certain features, however, that speak against this interpretation. First, the respondents were guaranteed anonymity in all surveys. This is likely to reduce the need to provide socially desirable answers. Second, there was a larger increase in the tolerant direction among those *con* participants who took part in the deliberation event compared to the *con* respondents in the control group. Since willingness to give desired answers can be expected to have an impact on both groups, this result suggests that social desirability cannot be the only explanation for opinion shifts and that deliberation had an impact on participants' viewpoints. Of course, one might argue that social desirability effect had to do with the fact that participants were at the university campus discussing in a group and being monitored by other participants and a moderator. This view does not, however, account for the fact that the *con* participants' opinions continued to change into a more liberal direction even after the event. Nevertheless, it should also be acknowledged that social desirability might be hard to separate from the actual effect of deliberation, and further, that social desirability is hard to distinguish from 'civilizing force of hypocrisy' (Elster 1998) which is likely to play a role in public deliberation.

The results of the experiment suggest that opinion polarization is not by any means an 'automatic' consequence of biases in group composition or, more precisely, in the initial dispositions of group members. One possible explanation for the absence of group polarization is the *ad hoc* nature of our experimental groups. Sunstein (2002, 180–181) points out the importance of "affective factors, identity and solidarity as factors which might increase or decrease group polarization". Affective ties can be expected to diminish dissent in groups. When people identify themselves as members of a group with a degree of solidarity, they are likely to reinforce the initial tendency prevailing in the group and, consequently, the group

becomes more extreme. Because the participants of our experiment were subject to random allocation into small-n groups, no affective ties should have been eminent. The *ad hoc* nature of the groups made it difficult for the participants to develop a within-group solidarity based on some common denominator (e.g. being pro- or anti-immigrant) in the given time frame of 4 h, especially since the participants were never informed about the composition of their group during the experiment.

In our experiment, all significant opinion changes were to the direction of more permissive attitudes toward immigration. The development of more permissive opinions in the mixed treatment may be explained by the established mechanism that exposure to differing political views “increases awareness of rationales for differing viewpoints and thus increases political tolerance” (Mutz 2006, 68). However, the increase in tolerance in *con* like-minded groups can still be regarded as somewhat surprising. Our experimental treatment which included two elements of what Mutz (2006, 61) calls the deliberative ‘package’, i.e. information and discussion procedures including rules and moderators, was not designed to disentangle their respective effects. Our finding that learning was not positively correlated with opinion changes suggests that the information provided in the briefing material, as such, did not contribute to the increased permissiveness of opinions.

The most likely explanation for our results is the nature of deliberation, after all. The results support the view that deliberation is different from other forms of talk and that discussion procedures can have a strong impact on outcomes. The ‘deliberative package’ seems to have an impact on how groups discuss and how opinions develop. This is, in fact, precisely what Sunstein with colleagues suggest when they discuss the results obtained in an experiment with like-minded groups without information, discussion rules or moderators: “(I)deological amplification is highly likely to occur as a result of political deliberation among the like-minded. It also suggests circumstances in which ideological amplification might be dampened or prevented. In particular, *interventions that involve external administrators or independent flows of information, might produce different kinds of shift and might serve to intensify or to dampen amplification*” (Schkade et al. 2010, 243, our italics).

Despite the initial ‘like-mindedness’ of the *con* enclave groups, there seems to have been a sufficient degree of disagreement to trigger deliberation where arguments were assessed by their merits. Information, discussion rules and moderators all encouraged the participants to evaluate arguments. Following Sunstein (2002, 180), polarization may have been avoided in the *con* like-minded groups because the participants who defended the prevailing tendency of the group were ‘particularly unpersuasive’, and the outliers (people with the most liberal immigration attitudes) were especially convincing. Furthermore, because individuals in the *con* like-minded groups did not know the composition of their group, they may have tried to argue in ways which would appeal also to people with conflicting viewpoints.

As pointed out by theorists of deliberative democracy, all arguments should *not* have an equal weight in processes of public reasoning. Most notably, reasonable arguments appealing to generalizable moral principles should be powerful, whereas arguments based on attitudes such as prejudice should be ‘laundered’ in the course of deliberation (see e.g. Goodin 1986). Liberal tendencies may have been reinforced

in the experiment because certain arguments against immigration were such that they could not be put forward or sustained in the deliberative process. The effects of preference laundering are likely to be particularly strong in an issue such as immigration where people's basic rights are at stake. Naturally, in order to gain more conclusive support for this kind of an assumption, the experiment should be replicated (a) on different types of issues and (b) in different political contexts.

Acknowledgments The authors would like to thank participants of several conferences and workshops for valuable comments and suggestions, most notably at the panel 'Inequalities in Political Knowledge: Individual and Contextual Determinants' at the IPSA World Congress in Madrid in July 2012, at the panel 'Racial Prejudice' at the WPSA Annual Meeting in Los Angeles in March 2013, at the panel 'The Effects of Interpersonal Discussion and Deliberation' at the MPSA Annual Meeting in Chicago in April 2013, and at the panel 'The Quality of Deliberation' at the ECPR General Conference in Bordeaux in September 2013. We are also grateful for the comments made by Prof. Lauri Karvonen, Prof. Peter Esaiasson, Prof. Marc Hooghe, the four anonymous referees as well as the Editors of Political Behavior. The experiment was funded by the Academy of Finland (Grant No. 251222) and the Centre of Excellence, 'Democracy: A Citizen Perspective' at Åbo Akademi University. The authors owe a debt of gratitude to the whole team behind the implementation of the experiment, especially Staffan Himmelroos, Marina Lindell, Lauri Rapeli, Maija Karjalainen and Kim Strandberg.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

Appendix 1

The main items measuring attitudes on immigration

1. Finland should take more immigrants. Do you think this is a bad or a good suggestion? [r]
2. Migration of foreigners into Finland should be restricted as long as there is unemployment in Finland. [r]
3. Do you think Finland will change into a better or a worse place to live when people from other countries move to Finland? (Questions 1–3 were presented on a scale from 0 to 10)
4. It is good for the Finnish economy that people from other countries move to Finland
5. Immigrants take away jobs from native Finns. [r]
6. Immigrants should have the same right to social security as Finns even if they are not Finnish citizens
7. The state and the municipalities use too much money to aid immigrants. [r]
8. Immigration poses a serious threat to our national originality
9. Everyone who wants to come to Finland to live and work should be allowed to do so
10. Immigration policy should primarily favor Christians instead of other religions. [r]
11. Generally speaking, immigrants adapt well into the Finnish society
12. I would be happy to have an immigrant as a co-worker
13. I would accept an immigrant as a family member
14. I would accept immigrants in my neighborhood. (Questions 3–14 were presented as a standard Likert scale with four values)

[r] = Reversed coding in the sum variable

Appendix 2

Knowledge gains within treatments (shares of correct answers)

Item	Like-minded				N = 119	Mixed				N = 88
	Before	After	Change	<i>p</i>		Before	After	Change	<i>p</i>	
1. Who decides on residence permits?	93	100	7	**	94	98	3			
2. The share of foreigners in Finland	61	83	23	***	55	73	18	**		
3. From which continent do most immigrants come to Finland?	63	88	25	***	48	90	42	***		
4. Who decides the size of the refugee quota?	40	66	27	***	41	64	23	***		
5. The size of the Finnish refugee quota	40	86	45	***	26	86	60	***		
6. The most common reason to apply for a residence permit	19	89	70	***	19	84	65	***		
Mean of correct answers (Info package items)	53	85	33	***	47	82	35	***		
7. The number of residents of foreign origin in (city to be inserted)	48	42	-6		39	36	-2			
8. Foreigners' share of sentenced crime	40	40	-1		39	40	1			
9. Unemployment among immigrants	24	24	1		23	28	6			
10. The amount of integration allowance in EUR	19	20	2		15	21	6			
Mean of correct answers (Items not in the info package)	33	32	-1		29	31	3			
11. The name of the foreign secretary	88	91	3		84	85	1			
12. Who are eligible voters in Finland?	72	71	-1		69	66	-3			
13. The definition of parliamentarianism	71	63	-8		61	52	-9			
14. The fourth largest party in the Parliament	56	58	3		56	55	-1			
15. The name of the EU (Lisbon) treaty	41	44	3		33	38	5			
Mean of correct answers (General political knowledge)	66	65	-1		61	59	-2			
Mean of correct answers (all 15 items)	52	64	13	***	47	61	14	***		

Significance (one-tailed *t* test) * < 0.05; ** < 0.01; *** < 0.001

Appendix 3

Socio-demographics of the participants and control group

	Participants (n = 207)	Control group (n = 369)
Gender		
Male	52.2	48.2
Female	47.8	51.8
Age		
18–34	22.2	16.3
35–49	16.4	15.2
50–64	27.5	33.6
65	33.8	35.0
Education		
Primary or Secondary	10.7	14.7
Vocational or Upper secondary	33.5	37.0
Polytechnic or Bachelor	34.0	35.1
At least Master's Degree	21.4	13.3

Appendix 4

Socio-demographics of the participants according to enclave and treatment

	Con like-minded (n = 42)	Con mixed (n = 44)	Pro like-minded (n = 77)	Pro mixed (n = 44)
Gender				
Male	50.0	47.7	44.2	52.3
Female	50.0	52.3	55.8	47.7
Age				
18–34	11.9	13.6	31.2	25.0
35–49	9.5	20.5	15.6	20.5
50–64	21.4	31.8	27.3	29.5
65	57.1	34.1	26.0	25.0
Education				
Primary or Secondary	14.6	20.5	6.5	4.5
Vocational or Upper secondary	34.1	47.7	26.0	31.8
Polytechnic or Bachelor	34.1	15.9	37.7	45.5
At least Master's Degree	14.6	15.9	29.9	18.2

References

- Andersen, V. N., & Hansen, K. M. (2007). How deliberation makes better citizens: The Danish deliberative poll on the euro. *European Journal of Political Research*, 46(4), 531–556.
- Asch, S. E. (1948). The doctrine of suggestion, prestige, and imitation in social psychology. *Psychological Review*, 55(5), 250–276.
- Delli Carpini, M. X., Cook, F. L., & Jacobs, L. R. (2004). Public deliberation, discursive participation, and citizen engagement. *Annual Review of Political Science*, 7, 315–344.
- Dryzek, J. (2000). *Deliberative democracy and beyond. Liberals, critics, contestations*. Oxford: Oxford University Press.
- Elster, J. (1998). Deliberation and constitution making. In J. Elster (Ed.), *Deliberative democracy* (pp. 97–122). Cambridge: Cambridge University Press.
- Farrar, C., Green, D. P., Green, J. E., Nickerson, D. W., & Shewfelt, S. (2009). Does discussion group composition affect policy preferences? Results from three randomized experiments. *Political Psychology*, 30(4), 615–647.
- Goodin, R. E. (1986). Laundering preferences. In J. Elster & A. Hylland (Eds.), *Foundations of social choice theory* (pp. 132–148). New York: Cambridge University Press.
- Goodin, R. E. (2000). Democratic deliberation within. *Philosophy & Public Affairs*, 29(1), 81–109.
- Goodin, R. E., & Dryzek, J. S. (2006). Deliberative impacts: The macro-political uptake of mini-publics. *Politics & Society*, 34(2), 219–244.
- Grönlund, K., Setälä, M., & Herne, K. (2010). Deliberation and civic virtue: Lessons from a citizen deliberation experiment. *European Political Science Review*, 2(1), 95–117.
- Isenberg, D. J. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology*, 50(6), 1141–1151.
- Jones, D. A. (2013). The polarizing effect of a partisan workplace. *PS: Political Science & Politics*, 46(1), 67–73.
- Karjalainen, M., & Rapeli, L. (2015). Who will not deliberate? Attrition in a multi-stage citizen deliberation experiment. *Quality & Quantity*, 49(1), 407–422.
- Karpowitz, C. F., Raphael, C., & Hammond, A. S. (2009). Deliberative democracy and inequality two cheers for enclave deliberation among the disempowered. *Politics & Society*, 37(4), 576–615.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480–498.
- Lee, E.-J. (2007). Deindividuation effects on group polarization in computer-mediated communication: The role of group identification, public-self-awareness, and perceived argument quality. *Journal of Communication*, 57(2), 385–403.
- Luskin, R. C., Fishkin, J. S., & Jowell, R. (2002). Considered opinions: deliberative polling in Britain. *British Journal of Political Science*, 32(3), 455–487.
- Mansbridge, J. (1994). Using power/fighting power. *Constellations*, 1(1), 53–73.
- Mansbridge, J., Bohman, J., Chambers, S., Estlund, S., Føllesdal, A., Fung, A., et al. (2010). The place of self-interest and the role of power in deliberative democracy. *The Journal of Political Philosophy*, 18(1), 64–100.
- Mendelberg, T., & Karpowitz, C. (2007). How people deliberate about justice. In S. Rosenberg (Ed.), *Can the people govern?* (pp. 101–129). Basingstoke: Palgrave Macmillan.
- Mercier, H., & Landmore, H. (2012). Reasoning is for arguing: Understanding the successes and failures of deliberation. *Political Psychology*, 33(2), 243–258.
- Mutz, D. C. (2006). *Hearing the other side. Deliberative versus participatory democracy*. Cambridge: Cambridge University Press.
- Mutz, D. C., & Mondak, J. J. (2006). The workplace as a context for cross-cutting political discourse. *Journal of Politics*, 68(1), 140–155.
- Noelle-Neumann, E. (1993). *The spiral of silence: Public opinion—Our social skin* (2nd ed.). Chicago: The University of Chicago Press.
- Sanders, D. (2012). The effects of deliberative polling in an EU-wide experiment: Five mechanisms in search for an explanation. *British Journal of Political Science*, 42(3), 617–640.
- Schkade, D., Sunstein, C. R., & Hastie, R. (2010). When deliberation produces extremism. *Critical Review: A Journal of Politics and Society*, 22(2–3), 227–252.
- Setälä, M., Grönlund, K., & Herne, K. (2010). Citizen deliberation on nuclear power: A comparison of two decision-making methods. *Political Studies*, 58(4), 688–714.

- Steenbergen, M. R., Bächtiger, A., Spörndli, M., & Steiner, J. (2003). Measuring political deliberation: A discourse quality index. *Comparative European Politics*, *1*, 21–48.
- Sunstein, C. (2002). The law of group polarization. *The Journal of Political Philosophy*, *10*(2), 175–195.
- Sunstein, C. (2007). *Republic.com 2.0*. Princeton, NJ: Princeton University Press.
- Sunstein, C. (2009). *Going to extremes. How like minds unite and divide*. Oxford: Oxford University Press.
- Thompson, D. F. (2008). Deliberative democratic theory and empirical political science. *Annual Review of Political Science*, *11*, 497–520.
- Torgerson, D. J., & Torgerson, C. J. (2008). *Designing randomised trials in health, education and the social sciences*. Basingstoke: Palgrave Macmillan.