EuComMeet
DEVELOPING ONLINE DELIBERATIVE SPACES

D 5.4 PROPOSAL FOR NEW PROTOCOLS ON POLARIZATION

WP 5 Polarization

Partner: Åbo Akademi University

Authors: Nanuli Silagadze & Kimmo Grönlund

Version: F

Date: 11.06.2024

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 959234

Project details and Deliverable information

Project details

Project Title	Developing Participatory Spaces using a Multi-stage, Multi-level, Multi-mode, Multi-lingual, Dynamic Deliberative approach (M4D2)
Project Type	Research and Innovation Action (RIA)
Project Acronym	EuComMeet
Grant Agreement No.	959234
Duration	36 months
Project Start Date	01/03/2021

Deliverable information

Status (F: final; D: draft; RD: revised draft):	Final
Planned delivery date	31/03/2024 (M37)
Actual delivery date	17/06/2024 (M40)
Dissemination level: (PU = Public; PP = Restricted to other program participants; RE = Restricted to a group specified by the consortium; CO = Confidential, only for members of the consortium)	PU
Type: Report, Website, Other, Ethics	Report

Document history and quality check

Document History

Version	Date (DD/MM/YYYY)	Created/Amended by	Changes
01	3/06/2024	ÅAU	First draft
02	11/06/2024	ÅAU	Final

Quality check review

Reviewer (s)	Main changes
UNISI	Not relevant

Table of Contents

List of Figures and Tables

List of abbreviations and definitions

Abbreviation	Definition
GA	Grant Agreement
DoA	Description of Action
EC	European Commission
H2020	Horizon 2020

Disclaimer

This deliverable may be subject to final acceptance by the European Commission. The results of these deliverables reflect only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Statement for open documents & Copyrights

Executive summary

This report synthesizes EuComMeet's experience regarding the issue of polarization. First, we explore the theoretical connection between deliberation and polarization and delve into the state of the art of automated moderation. Next, we present the effects of the deliberative experiment on levels of polarization in different group settings. Finally, we reflect on the lessons learned and offer suggestions for similar deliberative experiments in the future.

1. Introduction

Polarization has become one of the main concerns in a number of democracies with an increasing tendency of societies being divided into opposite camps. Accordingly, one of the key research questions within the EuComMeet project is how organized citizen deliberation can potentially alleviate polarization on highly divisive topics and to what extent an automated moderator can facilitate the reaching of mutual understanding among citizens.

The main experiment within the EuComMeet took place between May 8 – June 17, 2023. Nearly 400 European citizens from five different European Union countries (Italy, Ireland, France, Germany, and Poland) participated in the online deliberation. The deliberative process was divided into three different levels: local, consisting of individuals from the same city; national, involving citizens of the same country; and European, a blend of participants from previous rounds of discussions. Environment and climate change were the broad topics chosen for the deliberative event, further divided into subtopics: sustainable mobility, sustainable food consumption, and plastic pollution. During the period of deliberation, two methods were employed to engage citizens: synchronous sessions in the form of 90-minute video conferences and asynchronous sessions with a chat forum where text messages could be sent over the course of three days.

One of the unique features of the deliberative design was the use of automated moderation integrated into the deliberative platform. Citizens were assigned to a group with either automated or human moderators. This element is rather novel and has been applied previously only by the Stanford Online Deliberation Platform. The platform used for this project is developed by the EUComMeet partner NetHood and it is an open-source platform, hence, can be potentially used by various researchers and institutions for free (which is not the case with the Stanford platform).

2. DELIBERATION AND POLARIZATION

According to the literature, deliberative practices can help reduce polarization through informed and moderated dialogue and even strengthen anti-populist attitudes (Fishkin, 2018; Grönlund et al., 2015; Strandberg et al., 2019). Furthermore, deliberation can heel deep divisions in a society and promote understanding and learning about the other side (Niemeyer 2011). Studies show that deliberation can be employed as an effective method to reduce polarization on contentious issues – across countries and cultures. America in One Room – a recent national-wide field experiment with over 500 registered voters has shown that deliberation can dramatically reduce both substantive and affective polarization even in such a deeply divided society (Fishkin et al. 2021).

One of the central findings in recent scholarship has been that deliberation can counteract polarization even when the discussion takes place in like-minded groups. This can be explained by the presence of deliberative norms – facilitator and rules (Grönlund, Herne, and Setälä 2015). In one mini-public on immigration, opinions de-polarized rather than polarized even in the like-minded groups. More importantly, people with anti-immigrant attitudes became more tolerant even when they deliberate among people with the same anti-migrant attitudes on the issue (Grönlund, Herne, and Setälä 2015). Subsequent experiments confirmed this observation, demonstrating that de-polarization tendencies occur to an equal degree in both online and offline environments, making them an attractive option in the digital age (Strandberg, Himmelroos, and Grönlund 2019).

Deliberation is different from other forms of interaction and talk. Deliberative mini-publics are democratic innovations that are seen as a remedy for several problematic tendencies in today's democracies, including polarization and alienation from politics (Gutmann and Thompson 2004; Smith 2009). Deliberative mini-publics are defined as 'institutions in which a diverse body of citizens is selected randomly to reason together about an issue of public concern' (Smith and Setälä 2018, 300). Mini-publics can take several forms and they are being increasingly used in a variety of countries to complement representative democracy (Elstub and Escobar 2019; Grönlund, Bächtiger, and Setälä 2014). Furthermore, deliberation in a mini-public represents a 'structured discussion guided by a moderator and specific discussion rules' (Grönlund et al. 2021). Participants of a mini-public are usually selected through stratified random sampling, so that a range of demographic characteristics from a broader population are adequately represented, for example age, gender, ethnicity, geography, income, education. The idea is to assemble a microcosm of the public – 'a mini-public', with each citizen having an equal chance of being selected and thus, ensuring the legitimacy of the process.

Participants are commonly remunerated, discussions are facilitated, and experts are invited to provide balanced information (Escobar and Elstub 2017). The presence of a trained facilitator is seen as fundamental in shaping deliberation and making sure that marginalized voices get heard (Smith 2009). The moderator or facilitator ensures that the rules for deliberation are followed: they ask participants to speak up, to listen to others, to use arguments to persuade others of their viewpoints, to behave respectfully and overall try to make the deliberation experience pleasant for participants (Mansbridge et al. 2006).

In the discussion on polarization, it is important to differentiate between preference (aka. substantive or issue-based polarization) and affective polarization. Preference polarization refers to the increasing divide between individuals or groups in their political preferences and opinions. It refers to the widening gap in their views on political issues and politics in general. Affective polarization, on the other hand, refers to the growing emotional and psychological divide between individuals or groups in the political arena. This type of polarization is characterized by negative feelings and emotions towards those who hold different political views and/or belong to a different political party. In this case, animosity, distrust, and discrimination is displayed towards the out-group (the opposing party) whereas people show loyalty and favoritism towards their in-group (their own party). Affective polarization is measured by feeling thermometer ratings (Iyengar et al. 2019; Iyengar and Krupenkin 2018).

In today's world, discussions among like-minded groups have become dominant, especially in the online sphere. One of the dangers of this is group polarization, a phenomenon where a like-minded group of people becomes more extreme in their opinions as a result of discussion. This occurs partly due to the reinforcement of arguments in favor of the original position and the lack of alternative critical viewpoints (Sunstein, 2002, 2009). Fortunately, recent scholarship has found that deliberation can counteract polarization even when the discussion takes place in like-minded groups. This can be attributed to the nature of deliberative processes, which are usually well-thought-out and designed to mitigate the effects of groupthink. Trained moderators and specific rules for discussion play a crucial role (Fishkin & Luskin, 2004; Setälä et al., 2010). Participants are provided with a balanced set of information, both in favor of and against the topic at hand, often via experts and handouts.

In one mini-public on immigration, participants deliberated in both like-minded and mixed groups. Opinions de-polarized even in like-minded groups, with all participants improving their factual knowledge and correcting misperceptions. Individuals with anti-immigrant attitudes became more tolerant, even when discussing with others holding similar views. Overall, participants adopted more liberal attitudes towards immigration, with no one becoming more restrictive. Furthermore, the change in opinion was long-lasting –

participants remain more tolerant towards immigration even several months after the deliberative event (Grönlund, Herne, and Setälä 2015). Similarly, a mini-public on the use of nuclear power organized in Finland revealed that the participants moved towards centric positions rather than becoming more in favor or against the use of nuclear power (Himmelroos and Christensen 2014).

In summary, deliberation fosters deeper reflection, cultivates tolerance among citizens, increases political knowledge, trust, and perspective-taking among participants, and ultimately reduces polarization (Andersen and Hansen, 2007; Grönlund, Herne, and Setälä, 2015, 2017; Mercier and Landemore, 2012). Moreover, deliberation has a depolarizing effect in both online and offline settings (Strandberg, Himmelroos, and Grönlund, 2019), offering promise for policymakers seeking to engage citizens online and for social media platforms like Facebook or Twitter, which have faced criticism for exacerbating polarization. In the light of these effects, it is argued that deliberative mini-publics as a democratic innovation have potential to foster consensus-seeking in today's polarized societies (Herne, Christensen, and Grönlund 2019).

3. AUTOMATED MODERATION

In this section we introduce the state of the art in automated moderation, then describe the experience within the EuComMeet deliberative experiment followed by suggestions for future experiments.

3.1 STATE OF THE ART

One of the main incentives behind harnessing the technological developments and creating an online platform with automated moderation features is to replicate the offline deliberation. This approach aims to foster equitable and constructive participation while addressing scalability challenges by eliminating the need for recruiting and training neutral moderators. The Stanford Online Deliberation Platform is a pioneer self-moderating web-based platform that builds upon Deliberative Polling (Fishkin, 2011; Gelauff et al., 2023). Deliberative Polling is a popular framework for public deliberation which has been applied so far more than 110 times in over 34 countries resulting in policy impact in many cases. In the online version of this framework, deliberation is organized for a relatively large number of participants (typically 200–500 people) divided into smaller "rooms" (5–15 people each) with a shared agenda. Participants are surveyed before and after the deliberation in order to accurately measure the change in attitudes (Gelauff et al., 2023).

The Stanford Online Deliberation Platform is a pioneer self-moderating web-based platform that builds upon Deliberative Polling (Fishkin, 2011; Gelauff et al., 2023). Deliberative Polling is a popular framework for public

deliberation which has been applied so far more than 110 times in over 34 countries resulting in policy impact in many cases. In the online version of this framework, deliberation is organized for a relatively large number of participants (typically 200–500 people) divided into smaller "rooms" (5–15 people each) with a shared agenda. Participants are surveyed before and after the deliberation in order to accurately measure the change in attitudes (Gelauff et al., 2023).

3.2 EUCOMMEET EXPERIENCE

For the deliberative experiments, EuComMeet employed the Convivium platform, which blends two widely used free softwares, Nextcloud and Jitsi, with millions of users. This GDPR-compliant platform exemplifies collaboration between free software and public infrastructure, ensuring enhanced privacy, customization options, and economic sustainability. Convivium is developed by NetHood with the support of C.I.R.C.E. It is important to note that within the EuComMeet project, the approach to automated moderation was unique. It is more accurate to describe it as supervised moderation, as human moderators were behind the automated moderation features.

On the local level we experienced the highest amount of technical challenges compared to the national or European levels. This is understandable since although the platform was tested during a pilot phase, this was the first real-life exposure of the open-source AM platform to a high number of simultaneous users. Furthermore, some of the technical issues can be attributed to the complex design of the project as well as the lack of sufficient testing and training of moderators/participants.

On the national level, there were far fewer technical problems. Still, some challenges occurred: the platform was overloaded at times, the chat was not visible on smartphones, participants had problems accessing the attachments, and switching between chat and video was cumbersome. Similar to the local level discussions, a clear trend of informal moderation emerged. There was a tendency for one participant to take on a leadership role, reading out loud the bot's commands, giving floor to others, summarizing their statements, and ensuring a smooth flow of the conversation. The rest of the group seemed to appreciate the initiative and acted engaged.

The final debate on the EU level represented a unique technological and methodological challenge as participants had the opportunity to address local and European leaders. The participants were shown video contributions from politicians and administrators working within the European Committee of the Regions who would receive the results of citizens' deliberative engagement. Citizens could interact with each other and comment on communication materials they received from the politicians. To facilitate the discussion

between people from different countries, an automatic text message translation tool was introduced for the five distinct European languages of the participants. In sharp contrast to the local and national levels, there were no technical issues during the EU-level discussions apart from the translation.

In addition to synchronous sessions, asynchronous sessions were provided where participants could connect between real-time discussions. These asynchronous sessions were only available at the national and EU levels, not locally. There was no automated moderation in the asynchronous mode and participants were not informed if their offline discussions were moderated by a bot or by a human.

During the asynchronous deliberation, moderators posted instructions and tasks for participants every morning over three days, with reminders sent each afternoon. Depending on the group's response, moderators attempted to stimulate discussion, reintroducing questions if there was a lack of participation and engaging participants who hadn't contributed. For HM groups, moderators aimed to be more 'human' and personal, reassuring participants. For AM groups, they tried to be as neutral as possible and primarily using pre-recorded or default messages. Despite varying engagement levels, participants maintained polite and the interaction was equal throughout the sessions. However, discussions sometimes lacked depth, with participants sharing ideas without directly addressing questions or interacting meaningfully with others. Gender and age data were not provided, limiting insights into demographic dynamics.

4. Main Findings

There were 306 participants who participated in all three rounds of deliberation. They were assigned to three different groups based on their environmental stand[1]: cross-sectional, like-minded, and polarized. Furthermore, these groups were divided into three subgroups with different types of moderation: 107 of them deliberated with a human moderator (HM), 134 with an automated moderator (AM), and 65 had no moderator / social media environment (NM), serving as a control group (see table 1).

[1] Cross-sectional, like-minded and polarized groups were formed based on respondents' environmental position as measured during the preliminary survey with the following question: Some argue that "Environmental protection should take priority even at the cost of economic growth", others maintain that "Economic growth should take priority even at the cost of environmental protection", where would you locate yourself? You can choose any value between 0 and 10 to define your position. Cross-sectional groups consisted of participants a variety of viewpoints, reflecting the broader diversity of the population. In contrast to cross-sectional groups, like-minded groups are composed of individuals who share similar opinions. Polarized groups included participants with diametrically opposing environmental stances.

Table 1: Group composition

	Cross-sectional	***Like-minded***	***Polarized***	*Total*
HM	54	27	26	107
AM	64	38	32	134
NM	65	-	-	65
Total	183	65	58	306

There were four phases of the experiment:

1. Preliminary survey with invitation (T1)
2. Survey post-local deliberation (T2)
3. Survey post-national deliberation (T3)
4. Survey post-EU deliberation (T4)

4.1 AFFECTIVE POLARIZATION

To measure affective polarization participants were first asked which party they would absolutely not vote for (OUTPARTY) and then they were asked four questions on social distance about that same (least preferred) party which was coded as an index. The index showed acceptable levels of inter-item reliability, *a* = .95. These questions are:

1) How comfortable are you having close personal friends who are supporters of [OUTPARTY]?

2) How comfortable are you having neighbours on your street who are supporters of [OUTPARTY]?

3) How would you feel if you daughter or son married a supporter of [OUTPARTY]?

4) How would you feel about being a close co-worker with someone who is openly a supporter of the [OUTPARTY]?

The answers ranged from 1 Not at all comfortable, 2 Not too comfortable, 3 Somewhat comfortable, to 4 Extremely comfortable. 'Don't know' or 'prefer not to answer' responses were treated as missing values. Consequently, the lower the value, the higher the level of affective polarization. T1 represents the baseline attitudes towards the outgroup as recorded during the initial survey. T4 depicts the values after the EU-level deliberation.

Table 2: Development of Affective Polarization

	Pre (T1)	Post (T4)	T4-T2	S.E	t	p	N
All	2.59	2.59	0.00	0.04	-0.05	0.96	303
Cross-sec HM	2.68	2.58	-0.09	0.09	-1.05	0.30	54
Cross-sec AM	2.44	2.47	0.02	0.07	0.35	0.73	62
Cross-sec NM	2.70	2.64	-0.06	0.09	-0.66	0.52	64
Like-minded HM	2.51	2.59	0.08	0.12	0.67	0.51	27
Like-minded AM	2.72	2.70	-0.01	0.07	-0.19	0.85	38
Polarized HM	2.58	2.63	0.05	0.16	0.30	0.77	26
Polarized AM	2.46	2.58	0.12	0.10	1.24	0.23	32

Statistically significant at the * 0.05 level; ** 0.01 level; *** 0.001 level (two-tailed test)

As shown in Table 2, there were no significant changes or developments in regard to affective polarization. The mean difference for the whole sample is 0.00.

Table 3: Pearson's Correlation between affective polarization, age, and ideology

Variables	Pearson's Correlation	Number
Age & Affective polarization T1	0.04	303
Age & Affective polarization T4	0.06	303
Ideology & Affective polarization T1	0.20**	270
Ideology & Affective polarization T4	0.14*	270

Statistically significant at the * 0.05 level; ** 0.01 level; *** 0.001 level

It was found that age was positively correlated with social distance (affective polarization) scale as measured during the preliminary survey (T1)[2], r (301)= .04, p= .47 (see table 3). Considering the reverse coding of the variable affective polarization where the lowest value symbolizes the highest level of social distance, our findings suggest that those who are older display less affective polarization and those who are younger are more prone to be polarized towards the outgroup.

4.2 ISSUE POLARIZATION

To measure issue polarization another index was created which served as a proxy for people's attitudes towards climate change policies and their effectiveness, displaying acceptable levels of inter-item reliability, *a* = .75. The index consists of responses to the following three statements:

[2] Similar results were found when correlating age with social distance scale as measured after the final deliberation round, r (301) = .06, p = .34

1) The costs of the investments needed for a green transition are much higher than the costs of the damages due to climate change.
2) Climate change policies are counterproductive and even harmful to the economy, environment, and society more generally.
3) Climate change policies are ineffective, technological solutions will fix the problem in the future.

The responses ranged from 0 strongly disagree to 10 strongly agree. Accordingly, the higher the number – the more negative attitudes one has, the lower the number – the more positive. Accordingly, a negative mean change indicates an improvement in attitudes as a result of deliberation as compared to the initial survey.

Table 4: Development of issue polarization

	Pre (T1)	Post (T4)	T4-T2	S.E	t	p	N
All	4.54	4.33	-0.22	0.14	-1.52	0.13	306
Cross-sec HM	5.23	4.77	-0.46	0.26	-1.82	0.08	54
Cross-sec AM	4.77	4.99	0.22	0.53	0.43	0.67	64
Cross-sec NM	4.51	4.62	0.11	0.21	0.55	0.58	65
Like-minded HM	2.80	3.07	0.27	0.31	0.88	0.39	27
Like-minded AM	4.36	3.58	-0.78**	0.28	-2.84	0.01	38
Polarized HM	4.82	4.76	-0.06	0.33	-0.20	0.85	26
Polarized AM	4.49	3.27	-1.22***	0.32	-3.79	0.00	32

Statistically significant at the * 0.05 level; ** 0.01 level; *** 0.001 level (two-tailed test)

As shown in Table 4, overall there has been a slight improvement in attitudes for the whole sample (0.22 points). The biggest improvement occurred in the polarized automated moderation group, followed by the like-minded automated moderation group. These changes are statistically significant. It's noteworthy that both polarized groups experienced a positive shift in their opinions on climate change policies and their effectiveness. However, the fact that attitudes became slightly more negative in the like-minded group with a human moderator is rather peculiar and necessitates further investigation.

Table 5: Pearson's Correlation between issue polarization, age, and ideology

Variables	Pearson's Correlation	Number
Age & issue polarization T1	-0.16**	306
Age & Issue polarization T4	-0.05	306
Ideology & issue polarization T1	0.35**	272
Ideology & issue polarization T4	0.24**	272

Statistically significant at the * 0.05 level; ** 0.01 level; *** 0.001 level

Furthermore, a negative correlation between age and attitudes toward climate change policies was revealed as measured after the final deliberation round (T4), r (304) = -.05, p = .44 (see table 5). Considering the coding of the issue polarization scale, this implies that younger people have more negative attitudes toward climate change policies and their effectiveness than the older generation. Interestingly, the above-mentioned correlation after the final deliberation round was weaker than originally measured at T1, r (304) = -.16, p = .01. This suggests that the experience of deliberation influenced positively the youth's position. In general, the concern that people with with lower resources, particularly younger participants, women and less educated citizens adopt the opinions of the dominant groups in the deliberations is long standing (Sanders, 1997; Young, 2000). However, empirical research suggests that that sociodemographic inequalities cannot account for opinion changes in deliberative mini-publics (Hansen, 2004; Himmelroos & Christensen, 2014; Luskin et al., 2002).

4.3 ROLE OF GENDER, EDUCATION, AND IDEOLOGY

With gender and education, a few interesting trends were discovered. Firstly, there were more female participants (54%) as compared to males (46%). Men showed higher levels of affective polarization, whereas women were more critical of climate change policies and their effectiveness. When it comes to opinion/attitude change as a result of deliberation, women experienced a greater shift in their position on climate issues, the change was positive and statistically significant. Men barely changed their position, and the shift was statistically insignificant. As for affective polarization, the change was minor but in opposite directions: women became slightly more polarized and men – slightly less.

A peculiar trend concerning education and affective polarization emerged (see table 6). The more educated a participant was, the higher was their level of affective polarization. Moreover, people with basic and secondary education, became even more polarized as a result of deliberation, whereas people with tertiary education, became slightly less polarized. Results are statistically insignificant. The relationship between education and affective polarization is complex. Research shows that higher levels of education is not a warrant against affective polarization since education might lower affective polarization through weak authoritarian attitudes, but it can simultaneously raise it through engagement in politics (Han, 2022). Another study demonstrated that the better citizens fit into the socio-demographic profile of their party, the more affectively polarized they became (Harteveld, 2021b).

Table 6: Affective polarization across gender and education levels

	Pre (T1)	*Post (T4)*	*T4-T2*	*N*

Gender				
Male	2.46	2.49	0.03	139
Female	2.71	2.68	-0.03	164
Educational level				
Basic	3.03	2.97	-0.06	9
Secondary	2.66	2.62	-0.04	109
Tertiary	2.53	2.56	0.02	188

Statistically significant at the * 0.05 level; ** 0.01 level; *** 0.001 level (two-tailed test)

Table 7: Issue polarization across gender and education levels

	Pre (T1)	***Post (T4)***	***T4-T2***	***N***
Gender				
Male	4.43	4.44	0.01	140
Female	4.64	4.23	-0.40**	166
Educational level				
Basic	4.07	8.37	4.30	9
Secondary	4.94	4.46	-0.48*	109
Tertiary	4.34	4.06	-0.28*	188

Statistically significant at the * 0.05 level; ** 0.01 level; *** 0.001 level (two-tailed test)

As for issue polarization, a different trend occurred (see table 7). Respondents with secondary and tertiary education became significantly more positive towards the climate change policies and their effectiveness. Simultaneously, participants with basic education became very negative as a result of deliberation, even though they had the most positive outlook during T1. Although the number of respondents with basic education is low (9) and we cannot generalize the results, this is an interesting observation that requires further inquiry.

Another variable that is relevant to one's identity is the positioning on the left-right scale. This was measured by the following standard question: In politics people sometimes talk of "left" and "right". Where would you place yourself on the following scale, where 0 means the left and 10 means the right?".[3] A moderate positive correlation was revealed between the ideological positioning of the respondent and their attitudes toward climate change policies at T1, $r(270) = .35$, $p < .00$. (see table 5). This suggests that individuals with more right-wing political leanings tend to exhibit greater skepticism toward climate change policies and their effectiveness. This aligns with previous research indicating that individuals with right-leaning political inclinations tend to exhibit more climate change denial and opposition to climate policies compared to those

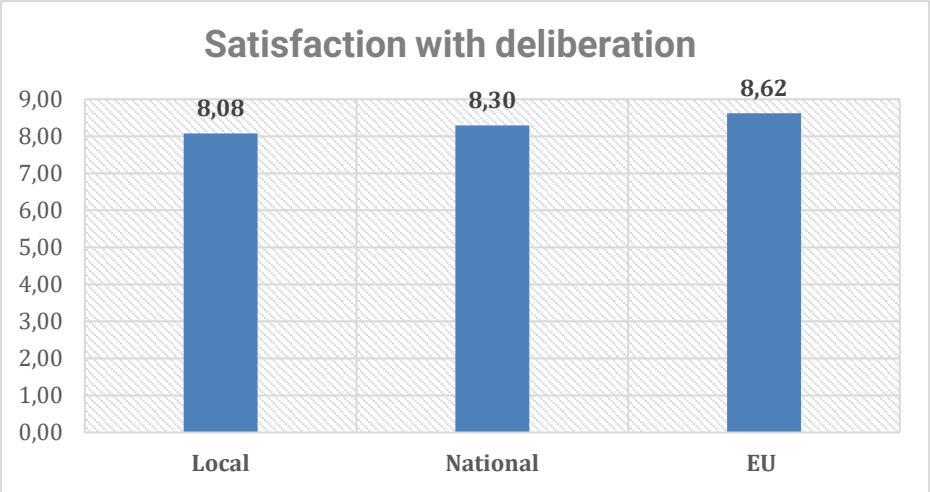
[3] Don't know / prefer not to answer / not able to locate on the scale were treated as missing values. It's noteworthy that around 5% of the respondents selected number 11 indicating that they are more right than 10, maybe eventually leaning towards far right. These responses with value 11 were recoded into 10.

leaning toward the left (Hornsey et al., 2016; Jylhä et al., 2020). It's noteworthy that this correlation became weaker at T4, r (270) = .24, p < .00 suggesting that deliberation was able to counteract the ideological positioning of individuals in relation to issue polarization on climate policies (table 6).

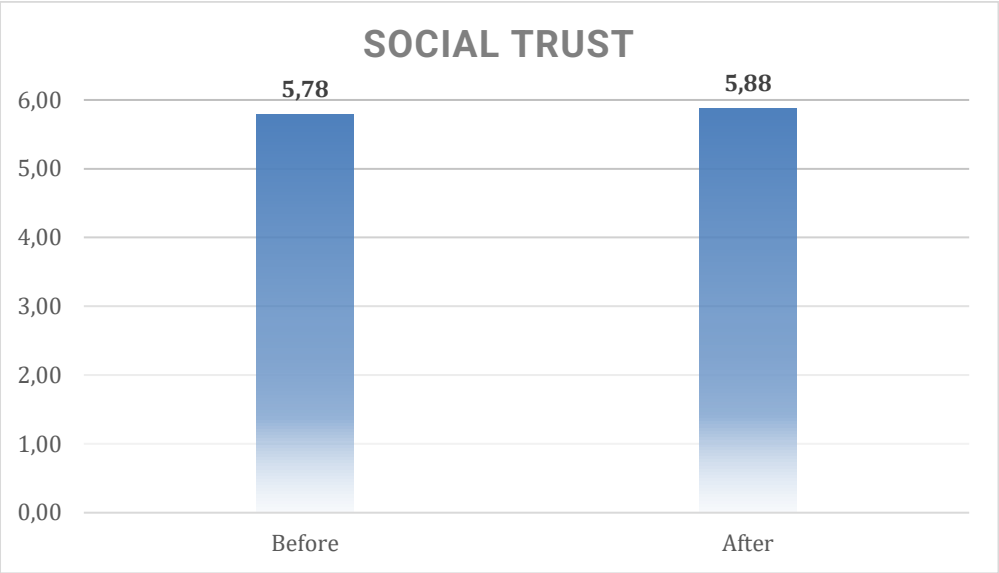
Ideological positioning certainly plays a role in affective polarization (Harteveld, 2021a). Our results confirm that there is a weak positive correlation between ideological leaning and social distance / affective polarization at T1, r (268) = .20, p < .00 (see table 2). This means that right-wing-leaning individuals experience lower levels of affective polarization. This contradicts the established literature according to which right-wing orientation is more conducive to affective polarization. For instance, dislike directed towards (far) right parties is on average more intense than dislike toward all other party families (Gidron et al., 2019). Furthermore, populist radical right not only emanate and receive higher levels of dislike, but their supporters are also uniquely negative about the supporters of mainstream parties and vice versa (Harteveld et al., 2022). Similar to issue polarization, the strength of correlation diminished as a result of deliberation at T4, r (268) = .14, p = .02 (table 2).

From previous research we know that deliberation represents a pleasant democratic experience for those who take part in it. As studies show, participants are satisfied with deliberation regardless of the level of disagreement in a group (Grönlund et al. 2021). Similarly, our results show that satisfaction with the deliberative event was quite high and increased further as participants moved through the levels (see graph 1). Furthermore, deliberation had a positive effect on the development of social trust (graph 2). Along the same lines, satisfaction with democracy increased across all the groups (graph 3). However, it is worth exploring why the highest level of satisfaction was noted among the groups without moderation. These findings represent good news from a democratic perspective – deliberative mini-publics create a safe environment where citizens can come together and deliberate on highly divisive political issues (for instance, immigration) without generating negative connotations. Furthermore, citizens become supportive of deliberation as a supplement to representative democracy after gaining first-hand experiences with mini-publics (Christensen, Himmelroos, and Grönlund 2017). This finding is in line with the classical work on participatory theory suggesting that participation generates an appetite for even more participation (Pateman 1970).

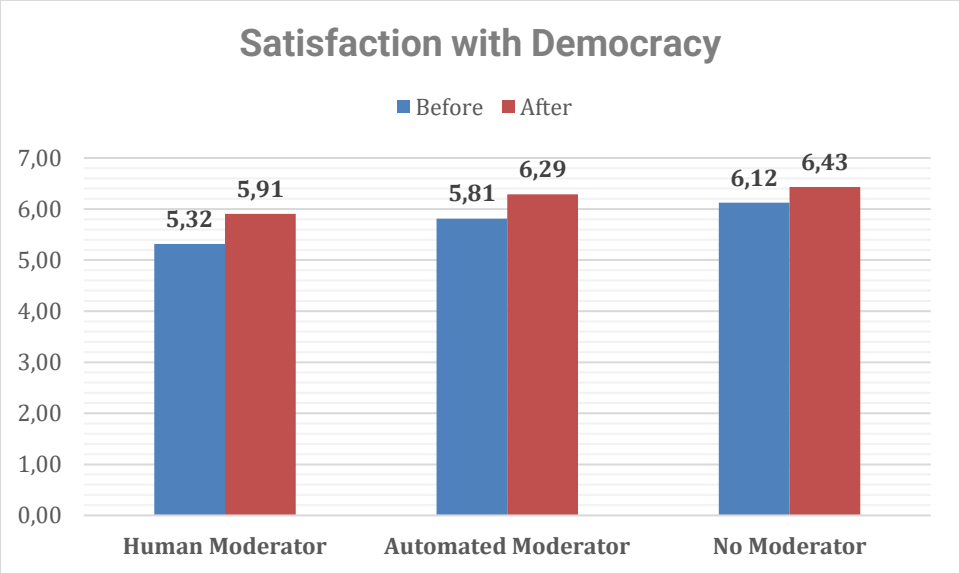
Graph 1: Satisfaction with the deliberative event



Graph 2: Development of social trust



Graph 3: Satisfaction with democracy across groups



5. LESSONS LEARNED

Overall, the design of the project has been experimental in nature and did not follow the classic deliberative model. Throughout the process, various technical issues disrupted the deliberation, with the local level being the most affected, while the national and EU levels experienced smoother operations. However, translation posed the greatest challenge at the EU level, as only English language translations functioned effectively, while other languages encountered difficulties. Alongside technical hurdles, time constraints posed another challenge, limiting the depth and quality of deliberation and output, as three sessions of 90 minutes each proved insufficient for discussing complex issues thoroughly. Additionally, high participant attrition led to the merging or cancellation of groups, resulting in some discussions involving only two individuals, significantly impeding meaningful debate.

From NetHood's standpoint, the most significant challenge of the project has been the intricate process of group management, particularly the merging and reshuffling of groups, which proved to be the most

demanding technical and procedural aspect. This challenge was closely linked to the overall project design. Additionally, connectivity issues, slow loading times, and inconsistent performance were addressed by the technical team, who conducted debugging based on extensive user feedback.

According to moderators, participation levels increased as the discussion progressed across different levels, with the highest level observed during the EU discussions. However, at the EU level, participants could only interact via chat with cameras on, which appeared somewhat awkward and limited the depth of debate, particularly due to issues with language translation. However, making direct comparisons is challenging due to variations in the number of participants and questions discussed across different levels.

The composition of the groups appeared to influence participants' satisfaction with the process. Those who transitioned from local discussions with fellow residents of the same city to national groups with representatives from different cities within the same country seemed to grasp the design and appreciate the experience of moving from local to national-level discussions. However, when groups were already mixed at the local level due to attrition, participants expressed frustration and a sense of redundancy with the national level deliberation, stating that they had already addressed these issues during the earlier discussions. Additionally, participants voiced frustration with the bot's functionality, noting that its instructions often did not align with the ongoing conversation. Many expressed dissatisfaction with the low level of interaction and the quality of discussions facilitated by the bot.

Regarding the quality of the deliberation, moderators noted a tendency among participants to reiterate their viewpoints across different stages, effectively becoming advocates for their preferred points or topics. Despite the progression through levels, participants appeared more focused on promoting their individual ideas rather than representing their city or country. This observation suggests that often participants remained ambassadors of their own ideas, only occasionally reflecting their group's discussions or exchanges. Regarding the asynchronous sessions, it was observed that participants largely shared similar opinions on the discussed topics, even in groups intended to gather polarized citizens. This raises concerns about the representativeness of the targeted demographics. It's possible that citizens skeptical of climate change were more difficult to recruit and engage in the conversation.

5.1 SUGGESTIONS FOR FUTURE PROTOCOLS

Below we summarized suggestions for improvements based on feedback from participants, moderators, and project partners for future deliberative events:

- Initiate the recruitment process earlier to mitigate high attrition rates among participants and avoid last-minute changes to group compositions, which added an additional layer of complexity to the project.
- Display participants by their names instead of identification numbers to facilitate communication. For moderators, it was uncomfortable to address individuals by the number instead of their names.
- Implement audio instructions or sound notifications for new instructions from the bot to ensure participants' attention during live conversations. Additionally, make instructions visually distinct and easier to read instead of having them appear as a regular message in the same chat where participants interact.
- Incorporate segments during deliberation to show images or provide brief explanations to enhance participant understanding of information material. Moderators got the impression that the participants did not read or understand the information material.
- Shorten the time allocated for asynchronous discussions to several hours per day and establish a minimum number of messages required from each participant to encourage maximum participation.
- Provide more extensive familiarization and training with the platform to make it more intuitive for the users to navigate Convivium. The team came to the conclusion that most of the issues were not technical, but were due to lack of familiarization and training. This also explains why we didn't experience similar challenges at national and EU level.
- Integrate additional tools into the platform, such as calendars, polling systems, whiteboards, and document collaboration, to enhance engagement, provided that there is sufficient time and effort dedicated to familiarization.
- Implement functions like a speaker queue or nudges (encouraging those who have not spoken in a while to share their viewpoint) and enable remote muting of participants for future experiments.
- A simpler project design would minimize technical challenges and allow for more straightforward scientific conclusions

It is imperative to acknowledge that this experiment was the first of its kind, and the platform is still in early development stage. Furthermore, AM and the Machine translation are still in early stages from a technical point of view. The optimization of platform performance is an ongoing endeavor, necessitating continuous feedback from users. The Stanford Online Deliberation Platform has facilitated so far 12 real deliberations

(with most of them being Deliberative Polls or minor variations) with a total of over 5,500 participants (Gelauff et al., 2023). As one of the participants noted: "Deliberation is similar to the UX design process, which must be repeated time and again". The same principle applies to the innovative open-source deliberation platform Convivium. At the same time, it is important to highlight that the Stanford platform is designed to support a well-established kind of deliberative interaction and does not allow for much variation. In contrast, Convivium can better adapt to and support experimental deliberative designs, such as the one used in this project. However, these designs are not intended to be repeated regularly and will not achieve the same degree of stability. An important suggestion remains to allocate enough time for proper testing and training when a new design is introduced.

6. Bibliography / References

Andersen, Vibeke, and Kasper M Hansen. 2007. "How Deliberation Makes Better Citizens: The Danish Deliberative Poll on the Euro." *European Journal of Political Research* 46: 531–556.

Christensen, Henrik Serup, Staffan Himmelroos, and Kimmo Grönlund. 2017. "Does Deliberation Breed an Appetite for Discursive Participation? Assessing the Impact of First-Hand Experience." *Political Studies* 65(1S): 64–83.

Elstub, Stephen, and Oliver Escobar, eds. 2019. *Handbook of Democratic Innovations and Governance*. Cheltenham: Edward Elgar.

Escobar, Oliver, and Stephen Elstub. 2017. "Forms of Mini-Publics: An Introduction to Deliberative Innovations in Democratic Practice." *newDemocracy*. https://www.newdemocracy.com.au/2017/05/08/forms-of-mini-publics/ (May 14, 2022).

Fishkin, James, and Robert C. Luskin. 2004. "Experimenting with Democratic Ideal: De-Liberative Polling and Public Opinion." *Paper presented at the Swiss Chair's Conference on Deliberation, European University Institute, Florence, Italy, May 21-22*.

Fishkin, J. S. (2011). When the People Speak: Deliberative Democracy and Public Consulation. Oxford University Press.

Fishkin, James S. 2018. Democracy When the People are Thinking: Revitalizing Our Politics through Public Deliberation *Democracy When the People Are Thinking: Revitalizing Our Politics through Public*

Deliberation. Oxford: Oxford University Press.

Fishkin, James, Alice Siu, Larry Diamond, and Norman Bradburn. 2021. "Is Deliberation an Antidote to Extreme Partisan Polarization? Reflections on 'America in One Room.'" *American Political Science Review* 115(4): 1464–81.

Gelauff, L., Nikolenko, L., Sakshuwong, S., Fishkin, J., Goel, A., Munagala, K., & Siu, A. (2023). ACHIEVING PARITY WITH HUMAN MODERATORS: A self-moderating platform for online deliberation1. In S. Boucher, C. A. Hallin, & L. Paulson (Eds.), The Routledge Handbook of Collective Intelligence for Democracy and Governance (pp. 202–221). Routledge.

Grönlund, Kimmo, André Bächtiger, and Maija Setälä, eds. 2014. *Deliberative Mini-Publics. Involving Citizens in the Democratic Process*. Colchester: ECPR Press.

Grönlund, Kimmo, Kaisa Herne, and Maija Setälä. 2015. "Does Enclave Deliberation Polarize Opinions?" *Political Behavior* 37(4): 995–1020.

Grönlund, Kimmo, Kaisa Herne, Kim Strandberg, and Peter Söderlund. 2021. "Disagreement and Deliberation: Evidence from Three Deliberative Mini-Publics." *Political Behavior* (online first): 1–23.

Grönlund, K., & Herne, K. (2023). Kansalaisparlamentin lehdistötilaisuus [Press-conference on the Citizens' Parliament]. https://www.kansalaisparlamentti.net/wp-content/uploads/2023/12/Lehdistotilaisuus_28.11.2023.pdf

Gutmann, Amy, and Dennis Thompson. 2004. *Why Deliberative Democracy?* Princeton: Princeton University Press.

Han, K. J. (2022). Education level and affective polarization: The mediation effects of psychological engagement in politics and authoritarian attitudes. Social Science Quarterly, 103(7), 1633–1646. https://onlinelibrary.wiley.com/doi/full/10.1111/ssqu.13228

Hansen, K. M. (2004). Deliberative Democracy and Opinion Formation. University Press of Southern Denmark.

Harteveld, E. (2021a). Fragmented foes: Affective polarization in the multiparty context of the Netherlands. Electoral Studies, 71, 102332.

Harteveld, E. (2021b). Ticking all the boxes? A comparative study of social sorting and affective polarization. Electoral Studies, 72, 102337.

Harteveld, E., Mendoza, P., & Rooduijn, M. (2022). Affective Polarization and the Populist Radical Right: Creating the Hating? Government and Opposition, 57(4), 703–727.

Herne, Kaisa, Henrik Serup Christensen, and Kimmo Grönlund. 2019. "The Influence of Political Knowledge on Opinion Polarization in Citizen Deliberation." *Political Research Exchange* 1(1): 1–23.

Himmelroos, Staffan, and Henrik Serup Christensen. 2014. "Deliberation and Opinion Change: Evidence from a Deliberative Mini-Public in Finland." *Scandinavian Political Studies* 37(1): 41–60.

Iyengar, Shanto et al. 2019. "The Origins and Consequences of Affective Polarization in the United States." *Annual Review of Political Science* 22: 129–46.

Iyengar, Shanto, and Masha Krupenkin. 2018. "The Strengthening of Partisan Affect." *Political* Psychology 39: 201–18.

Luskin, R. C., Fishkin, J. S., & Jowell, R. (2002). Considered Opinions: Deliberative polling in Britain. British Journal of Political Science, 32(2), 455–487.

Mansbridge, Jane, Janette Hartz-Karp, Matthew Amengual, and John Gastil. 2006. "Norms of Deliberation: An Inductive Study." *Journal of Deliberative Democracy* 2(1): Article 7.

Mercier, Hugo, and Hélène Landemore. 2012. "Reasoning Is for Arguing: Understanding the Successes and Failures of Deliberation." *Political Psychology* 33(2): 243–258.

Niemeyer, Simon. 2011. "The Emancipatory Effect of Deliberation: Empirical Lessons from Mini-Publics:" *Politics & Society* 39(1): 103–40.

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

Sanders, L. M. (1997). Against Deliberation. Political Theory, 25(3), 347–376.

Setälä, Maija, Kimmo Grönlund, and Kaisa Herne. 2010. "Citizen Deliberation on Nuclear Power: A Comparison of Two Decision-Making Methods:" *Political Studies* 58(4): 688–714.

Smith, Graham. 2009. *Democratic Innovation: Designing Institutions for Citizen Participation*. Cambridge: Cambridge University Press.

Smith, Graham, and Maija Setälä. 2018. "Mini-Publics and Deliberative Democracy." In *The Oxford Handbook of Deliberative Democracy*, eds. Andre Bächtiger, John S. Dryzek, Jane Mansbridge, and Mark E. Warren. Oxford: Oxford University Press, 300–314.

Strandberg, Kim, Staffan Himmelroos, and Kimmo Grönlund. 2019. "Do Discussions in Like-Minded Groups Necessarily Lead to More Extreme Opinions? Deliberative Democracy and Group Polarization." *International Political Science Review* 40(1): 41–57.

Sunstein, Cass R. 2002. "The Law of Group Polarization." *Journal of Political Philosophy* 10(2): 175–95.

Sunstein, Cass R. 2009. *Going to Extremes. How Like Minds Unite and Divide*. Oxford: Oxford University Press.

Young, I. M. (2000). Inclusion and Democracy. Oxford University Press.