Representative Personalities?

A large-scale field experiment on a citizen assembly in Austria

Ulrike Bechtold Katharina Gangl Marcela Ibanez
Gerhard Riener Marcel Seifert

Citizen Assemblies are a popular direct democratic instrument but are often criticized for not being representative of the population. A Citizen Assembly consists of randomly invited citizens who can self-select into participation. We present a large-scale field experiment in Austria, showing that under standard invitation rules participants are indeed not representative in terms of an important trait: Locus of control. We devise an intervention where we vary the invitation letter in two treatments, deviating to emphasize either personal experiences ("Experience" treatment) or the importance of personal perspectives by stressing the need of different opinions hence rendering those more valuable ("voice" treatment), rather than the requirement to bring own ideas. The voice treatment increases the number of applicants and makes the assembly more representative in terms of Locus of Control.

1 Deutsche Zusammenfassung

Bürger:innen-Räte sind ein beliebtes direktdemokratisches Instrument, werden jedoch oft dafür kritisiert, nicht repräsentativ für die Bevölkerung zu sein. Ein Bürger:innen-Rat besteht aus zufällig eingeladenen Bürger:innen, die freiwillig teilnehmen können. Wir präsentieren ein großes Feldexperiment in Österreich, das zeigt, dass die Teilnehmer:innen unter den üblichen Einladungsregeln tatsächlich nicht in Bezug auf ein wichtiges Persönlichkeitsmerkmal repräsentativ sind: der Kontrollüberzeugung. Wir entwickeln eine Intervention, bei der wir den Einladungsbrief in zwei verschiedenen Versionen weiterentwickeln, die entweder persönliche Erfahrungen betonen oder die Wichtigkeit die persönliche Perspektive einzubringen . Die Einladungsbriefe die betonen, wie wichtig es ist, dass persönliche Perspektiven und eigene Stimme eingebracht wird kann mehr Menschen zur Teilnahme mobilisieren und den Bürger:innen-Rat repräsentativer im Sinne der Kontrollüberzeugung machen.

2 Introduction

Democracy operates as a mechanism for resolving conflicts peacefully, underscoring the need to master the art of productive disagreement. Yet, the development of such critical skills is predominantly limited to elite educational institutions, leaving many without the tools necessary for effective participation. In these settings, opportunities to contribute ideas are not equitably distributed, often leading to a suboptimal gathering of perspectives. Citizen assemblies, as a form of a mini-populus that Dahl (1989) defined as an "attentive public that represents the informed judgement of the demos" Lacelle-Webster and Warren (2021) recognize this suggest to incorporate a wide array of societal views into the democratic process to achieve representation of opinions within society.[1] However, achieving this level of representativeness presents ongoing challenges. Administrative efforts to ensure balanced participation are frequently compromised by self-selection biases and the difficulties involved in effectively engaging a diverse citizenry. It is essential, therefore, to develop strategies that not only reach out to but actively engage a broad spectrum of citizens, thereby enhancing the legitimacy and inclusiveness of democratic decisions.

Citizen assemblies have recently emerged as a popular method to promote democratic engagement. Nevertheless, the true representativeness of these assemblies often comes under scrutiny. Previously hailed as effective instrument for direct democracy in Ireland, recent Irish Citizen Assemblies played a major role in the resignation of the Irish Prime Minister after the assembly's recommended bills did not align with mainstream political opinions and were subsequently rejected in a public vote which contributed to his decision to resign (Carroll 2024). Moreover, political opponents and also the media challenged legitimacy of decisions within such assemblies as they doubted its representativeness.

A common strategy to target this criticism, also applied in the Irish case, was setting quotas based on observable characteristics like gender, race, or age, it falls short in addressing equally crucial but less observable traits such as personality or locus of control. These traits are challenging to quantify and, at times, controversial to target directly. An effective alternative is the design of the invitation process, which can influence who feels encouraged to participate. The wording of the invitation can help alleviate fears and reservations about expressing opinions or sharing personal experiences, thus fostering a turnout that better represents the community. This strategy's success also depends on the pre-testing of invitations, as they not only serve as the first point of contact but also set the initial tone of the assembly and can vary in effectiveness based on the topic at hand.

[1] One might think here also in incorporating different discourses within a society, that including individuals who are not self affected but are speaking or representing a certain view within society and claiming to speak for a larger group, although Dryzek and Niemeyer (2008) argue that this can have legitimacy and importance in deliberative democratic institutions, we do not take this route, as it raises questions of conceptualization of discourses in the light of the problem of aggregation of individual preferences. We therefore stay strictly on the individual level of analysis.

The challenge of engaging a diverse group of citizens in assemblies likelystarts at the invitation stage, where the framing of the invitation significantly influences potential participants' perceptions of their value and potential contributions, that previous qualitative research has identified as a potential driver for non-participation in assemblies (Jacquet 2017) that finds its theoretical underpinning in the ideas of Osborne, Rosenthal, and Turner (2000). While Hibbing and Theiss-Morse (2002) argue that it is reasonable that most Americans do not want more deliberative democracy and it is efficient as "the average citizen is ill equipped to discharge the duties that deliberative theorists would assign to them" (Neblo et al. 2010, 568), it does not seem normatively justified to ignore those opinions. It is therefore important to understand how the presentation of the assembly can address fears and apprehensions about participating, which can encourage a turnout that more accurately mirrors the community. Recognizing the scepticism with which government invitations are often viewed, especially by those whose perspectives are most vital yet hardest to reach, introduces additional complexity.

This study aims to examine how different formats and contents of invitation letters affect citizens' willingness to participate in a citizen assembly, seeking to identify effective strategies for enhancing engagement and representativeness. The experiment includes 14,999 participants, a representative (according to observables of age, gender, and education), randomly selected sample of an entire nation. We assigned these subjects to three different treatments, each receiving a different invitation letter. The first treatment, "Ideas" (the standard invitation letter), emphasizes the importance of the participants' ideas and solutions. The second treatment, "Experience", highlights the significance of the participants' experiences and problems. The third treatment, "Voice", underscores the importance of the participants' personal perspectives. This directly addresses the perception of the "duties discharged" to the average citizen.

We focus on two sets of main outcomes: the willingness to participate in the assembly and the actual participation and the distribution of internal and external locus of control among the participants. Locus of control is a psychological construct that captures the extent to which individuals believe they can influence outcomes through their actions. It is a key determinant of economic and social outcomes, influencing individuals' decisions and behaviours in various domains and is " 'a generalised attitude, belief or expectancy regarding the nature of the causal relationship between one's own behaviour and its consequences' Rotter (1966). Those believing that life's outcomes are due to their own efforts have an internal locus (sense) of control, while those believing that outcomes are due to external factors (e.g. luck) have an external locus (sense) of control (Gatz and Karel 1993)." (cited from Cobb-Clark and Schurer (2013).)

We provide a first step in the direction of influencing representativeness on unobservable through minimal-cost interventions. Having locus of control as our first candidate variable of typically unobserved traits where there is a large gap between. What does true representativeness mean? Political affiliation, people who normally do not speak up in public, so inclusiveness is an important issue.

Our findings are in line with our hypothesis, derived from models of political participation Osborne, Rosenthal, and Turner (2000), Lohmann (1994), Andor et al. (n.d.). First we observe that there is a substantial amount of selection with respect to the locus of control. Compared to an arguable more representative sample of the Austrian population, participants in the assembly have a higher locus of control across all treatments. We find that wording of the invitation letter has a significant effect on the willingness to participate in the assembly. The *Voice* treatment has the highest effect on the willingness to participate in the assembly. The *Experiences* treatment has a smaller effect on the willingness to participate in the assembly all compared to our baseline, the standard *Ideas* invitation.

We can improve on the diversity of participants, especially young cohort men with easily scalable methods. Previous interventions to increase citizen participation in deliberative democratic activities may have found similar sized effects of around 14%, however those activities— engaging with members of the US congress or providing substantial financial incentives— as examined by Neblo et al. (2010) are not easily scalable and might lead to different selection effects. This is important as we possibly need to attract more citizens who feel most excluded from political processes, trying to close a niche by providing them with a voice that otherwise could be filled by parties with anti-democratic tendencies.

2.1 Literature

Our research advances the discussion on representativeness in political participation by critically evaluating the methodologies and assumptions in the existing literature. Previous studies primarily focused on visible demographics and quota systems in political assemblies, such as Chattopadhyay and Duflo (2004) 's examination of mandated representation of women in India. However, these studies often overlook how recruitment methods, like the wording of invitation letters, impact diversity in political participation not only in observables but in attitudinal attributes and personality traits, that are correlated with the willingness to participate (Jennstål 2018). Our study extends these insights by investigating how these methods influence both demographic and psychographic representativeness.

Bächtiger et al. (2018) and Lacelle-Webster and Warren (2021) highlight the role of demographic representativeness in enhancing the diversity of perspectives and improving the quality of deliberations, referencing foundational works by Dryzek and Niemeyer (2008) and Landemore (2017). They suggest that including a wide range of life experiences and values can improve the deliberation process. We argue that the recruitment phase is critical in shaping the composition and output of these assemblies.

The literature acknowledges the risk of assemblies attracting individuals with extreme or outlier preferences. While Warran (2015) suggests that stratified random sampling could mitigate these tendencies, Lacelle-Webster and Warren (2021) note the management of such dynamics remains poorly defined, without clear benchmarks for assessing outcomes. Our research fills

this gap by examining how different recruitment strategies impact representativeness, moving beyond simple demographic considerations.

We also contribute to the literature of psychological traits and political participation (A. S. Gerber et al. 2011), especially the locus of control, correlate with political participation and affect the outcomes of citizen assemblies. The closest research to this respect is Jennstål (2018), who particularly assessed the role of the BIG5 personality traits in the role of invitations to assemblies. Andor et al. (n.d.) shows that individuals with a higher internal locus of control are more likely to engage actively in civic activities, such as voting in significant elections. We propose that these psychological traits could predict participation levels and effectiveness in political assemblies.

We also revisit models of political participation that emphasize the perceived effectiveness of influencing political outcomes (Lohmann (1994); Osborne, Rosenthal, and Turner (2000)). Our findings suggest that personal efficacy beliefs, such as the locus of control, critically influence individuals' decisions to engage in political processes, expanding the framework suggested by Andor et al. (n.d.).

We also touch on the literature of the effect of political deliberation. Ban, Jha, and Rao (2012), Fishkin et al. (2024)

Finally, our study contributes to the debate on external validity in experimental research, focusing on citizen assemblies intended to reflect broader societal deliberations. Similar to field experiments, achieving representative samples in such settings is challenging due to self-selection and stratification issues. Drawing on scaling literature in field experiments Al-Ubaydli et al. (2017), Vivalt (2020), Maniadis, Tufano, and List (2014), Riener, Schneider, and Wagner (n.d.), we show that enhancing representativeness in political experiments requires complex strategies beyond demographic stratification, affecting both the reliability and applicability of the findings in societal contexts. This approach not only addresses design concerns but also informs the ongoing discussion on inclusivity and representativeness in political participation.

3 Conceptual framework

The participation in a citizen assembly that gathers information can be seen as a public good problem similar to the participation in voting, with the crucial difference that the marginal cost of participation is higher and the potential influence larger. A individual motivation to contribute can be seen as warm glow (Andreoni 1990) or as a strategic move influence policy makers information on what may be expected over the median voters preferences (Lohmann 1993). Independent of this motivation, and augmenting it with a notion of perceived ability to contribute, i.e., the perceived individual multiplier to the contribution to the public good.. This perceived ability we assume to be a function $a(\cdot)$ increasing in the personal locus of control l_i and decreasing in their belief over the difficulty of the requested requirements for useful contribution r_i. We follow Andor et al. (n.d.) who take the following interpretation of

the locus of control (LoC) from Paulhus (1983): LoC can be understood as a control belief over individual outcomes ("Can I solve the problem?") as well as a control belief over socio-political outcomes ("Can society solve the problem?"), so this can be seen as some form of effectiveness to realise the idea. The requested requirements r_i are the individual's perceived usefulness of their ideas given the stated requirements. Moreover an individual is endowed with time, money, or other resources that we denote by m_i . This leads us to the following utility function for individual i:

$$U_i(g_i, G_{-i}) = m_i + a(l_i, r_i)b_i G(g_i, G_{-i}) + a(l_i, r_i)w_i g_i - cg_i$$

,

where $a(\cdot)b_i$ represents the agent's LOC and ability weighted utility gain from a one-unit increase in the total amount of the public good provided and $a(\cdot)w_i$ captures the ability weighted "warm-glow" utility. Under this setup, an individual will participate if $a_i(b_i+w_i) \geq c$, i.e., if the perceived benefit of the contribution exceeds its cost. Hence, the propensity for the individual to participate in an assembly is increasing in (external) LoC and decreasing the the requested requirement.

This allows for the two types of our main research hypothesis, the first set on the treatments changing the communicated requirements and the second one on LoC. Ceteris paribus, we assume that a lower requirement increases the perceived ability to contribute and hence the likelihood of participation. Moreover, we expect a positive association between LoC beliefs and participation.

Moreover, we hypothesize that the cross-partial derivative of a with respect to l_i and r_i is positive, i.e., the more the individual believes in their ability to contribute and the more demanding the requirements, the more likely they are to participate. Therefore, we expect that the average LoC belief is lower in the treatment with the least demanding requirements.

Conjecture 3.1 (Treatment comparison of participation)**.** (a) *The less demanding the invitation, the more subjects will participate in the citizens' councils. Specifically, invitations emphasizing "different views" will attract more citizens than those emphasizing "solution development ideas" or the inclusion of "personal experiences".*

(b) *These effects are particularly pronounced among individuals without high school degree.*

4 Experimental setup

In this study we recruit citizens for a citizen assembly in Austria using different methods of invitation. The aim is to address and encourage subjects who feel politically disenfranchised to participate. To do so, we vary crucial texts in the invitation letter. In this citizen assembly that took place in autumn 2023, measures for coping with future crises were developed by

participants, moderated by professional moderators. In each province the 40 citizens spent the day in an iterative mix of plenary input and and discussion rounds in small groups of about 7 persons. We will call the discussion round *assembly*. The citizens' councils were initiated by the federal government to deal with the Corona pandemic and are scientifically accompanied by the Austrian Academy of Sciences (OeAW) and the Institute for Advanced Studies, Vienna (IHS). 14,999 citizens invited, block stratified over the municipalities. They received an invitation from the Office of National Statistics, Austria (Statistik Austria). The invitation was sent out in three different versions, which were randomly assigned to the citizens. The three versions differed in the emphasis on the citizens' own experiences, the development of solutions, and the representation of different views. The study was registered at osf.io (Riener and Gangl 2023). The study was approved by the ethics committee of the Institute for Advanced Studies, Vienna.

4.1 Structure of the assembly

The series of nine assemblies organized across the country's federal states engaged 32 to 40 participants each, further divided into smaller groups of approximately eight members, leading to a total potential engagement of up to 360 individuals. The assemblies convened for a full day, from 9 am to 5 pm, focusing on evaluating the policies implemented during the pandemic and identifying areas for improvement in anticipation of future crises. This objective guided the discussions, providing a structured framework for participants to assess the effectiveness of various strategies and to propose recommendations for more resilient public health and crisis management systems.

Following the model of previous assemblies, such as those addressing the climate crisis held throughout the European Union [CITATION], the structure facilitated a mix of smaller group discussions and larger plenary sessions. This format allowed for in-depth structured discussion of topics in the smaller groups, while the plenary sessions served to consolidate insights and share findings across the entire assembly. These gatherings took place on weekends, predominantly on Saturdays, with one exception on a Sunday, in centrally located and easily accessible public buildings within the federal capital cities. This choice of venue and timing was strategic, aimed at maximizing participation and ensuring the process was as inclusive as possible.

By aligning with the methodologies and operational dynamics of prior assemblies, such as the Climate Crisis assemblies, these sessions aimed to harness collective intelligence in tackling complex issues. The discussions not only focused on critiquing past actions but also on co-creating future-oriented solutions, incorporating the experiences and perspectives of participants.

5 Experimental design

The experiment consisted of an invitation stage and the actual citizen assemblies in the capitals of the nine Austrian states. The treatments were applied at the invitation stage and the covariates used were administrative data collected before the application of the treatments and the data collected during and after the citizen assembly. Moreover, independent of the assembly, we conducted a survey to compare the general population with the participants of the assembly.

5.1 Invitation stage

Three treatments were stratified over age group and gender and blocked by federal province where in each province roughly equally many invitations were sent out (See Table XXX in the Appendix [TBD]), despite their differences in size. This lead to an overweighting of small states, that we correct for in the analysis using state population weights from July 2023 to estimate a nation wide Average Treatment Effect. Moreover, we conducted a reference questionnaire to compare the outcome variables for the self-selected assembly participants with a far less selected sample. While the size of the Austrian states varies substantially, in each state the same number of of citizens were invited to the assembly.

The treatments were embedded in the invitation letter sent out in two batches on August 9 and August 22, 2023 by Statistik Austria (see facsimile in German in the Appendix and a translation of a generic letter). We varied parts of the text on the first page of the letter, in a box that made the purpose of the assembly salient. The three treatments varied along the cognitive demand and inclusivity of opinions of what will be asked for within the assembly.

Clearly, our treatment variation can only effect those who effectively opened and read the letter. We do not have a measure of how many subjects opened the letter, but we can be assured that the treatment was applied as intended, as the letter was sent out by a public authority and the treatment was the only content of the letter. These were the treatments:

5.1.1 Voice

This invitation just focused on that the voice of people will be heard. It was formulated in the direct invitation as that the views are different and personal perspectives will be considered. It is important to also give a voice to those people who have not been heard much before. This ensures that all opinions are expressed in a balanced way and that you are also heard.

Exact wording

5.1.2 Experience

Sharing Experiences 'Experiences and problems are different'. Your experiences and problems matter. You have the opportunity to work together to write recommendations for the decision-makers:inside. This ensures that all opinions are balanced and that you are heard.

Exact wording

5.1.3 Ideas

This treatment is the most cognitive demanding, where the subjects' ideas and proposed solutions and opportunities for improvement were explicitly mentioned. You will have the opportunity to collaboratively write recommendations for the decision makers. This ensures that all opinions are balanced and that you are heard. These type of frame is typically used and recommended for invitation letters (CITATIONS).

Exact wording

In the subsequent analysis, we will use the least cognitive demanding, *Voice*, as baseline outcome and compare it to the treatments *Ideas* and *Experience*.

5.1.4 Data and compliance

We have administrative data on province of residence, sex (binary male/female), age group (18-35, 36-59, 60-89) and education level (binary, high school/no high school). Treatments were applied using personally addressed physical mail, protected by criminal law §118 StGB (Austrian Criminal Code). No other person was allowed to stand as substitute for the invited. Our outcome variable is the stated willingness to participate in the assembly, as recorded on a reply card sent back to Statistik Austria and actual participation as recorded by the enumerators at the locations of the assembly. This can be seen as complete monitoring of the treatment application procedure. No failure of protocol has been reported by the responsible entities. We also consider the probability of knowing other participants and communicated about this as low. Hence, we consider the stable unit treatment value assumption (SUTVA) (Cox 1958) to be fulfilled.

5.2 Reference Survey

We administered both online and telephone surveys to a sample of the Austrian population that reflected the demographics of the assembly survey. Our target was 2,000 participants, aged 18 to 59, representing a balanced mix of age, gender, and state. Of these, 1,800 were surveyed online, while 200 were reached by phone. This survey mirrored the survey administered to the

assembly participants and was conducted in the same time frame, allowing a direct comparison between assembly participants and the general population in Austria.

5.3 Citizen assembly stage

From the sample of those who actually show-up at the citizen assembly. A detailed questionnaire can be found in the appendix. The sample of those who come to the citizen assembly will be filling a questionnaire three times: (Wave 1) at the beginning of the assembly (Wave 2) at the end of the assembly (Wave 3) two weeks after the assembly. The majority of questions will be the same at all three points in time

We collected measures of the Locus of control at Wave 1 and the Reference Survey. The locus of control, as conceptualized Rotter (1966) in 1954, refers to an individual's belief system regarding the causes of his or her experiences and the factors to which that person attributes success or failure. This construct distinguishes between two types of control: internal locus of control, where individuals believe they can influence outcomes through their actions, and external locus of control, where individuals feel that their outcomes are determined by external forces beyond their control. We use the scale by (Kovaleva et al. 2012). It consists of a series of statements to which respondents indicate their level of agreement, assessing the extent to which they attribute control internally or externally. This scale is instrumental in understanding behavioirs in various fields, including economics, by linking individuals' perceptions of control to their economic decisions and behaviours, which will be the focus of a companion paper.

The variables we collected across all three waves were trust in institutions, polarization on Covid-19 protective measures, and concern about polarization that are important indicators to understand public sentiment and behaviour. These measures take from previous research, are tailored to capture specific nuances relevant to the Austrian post-Covid-19 context.

Our trust in Institutions measures assesses individuals' confidence in the various institutions that were important in handling the crisis. We asked specifically about trust in public service, private media, social media, and local media. Trust in public and private research institutions. Trust in the EU, Austrian federal government, state government, and local politics. Trust in people at national, state, and neighbourhood levels. We applied a 5-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree".

For the Polarization on we use the scale by (Boxell et al. 2022), this variable measures the extent to which opinions on Covid-19 related measures (such as lock-downs, mask mandates, and vaccination campaigns) are divided along ideological or partisan lines. Unlike the original study by (Boxell et al. 2022), which may have focused on general polarization in digital media consumption or its impact on political polarization, our study specifically addresses polarization regarding Corona measures. This could involve assessing agreement or disagreement with these measures, perceived effectiveness, and compliance, potentially identifying factors that contribute to polarized views (e.g., political affiliation, media consumption habits, trust in science).

Measuring a general concern about Polarization we draw on the framework of (Doherty, Kiley, and Asheer 2019), this measure evaluates how much individuals are worried about the effects of polarization, specifically in the context of the COVID-19 pandemic. This could involve concern over societal division, the erosion of public discourse, or the impact of polarization on the effectiveness of public health responses. This measure acknowledges that polarization itself can be a source of concern, influencing individuals' mental health, sense of community, and perspectives on future societal challenges.

These measures, by examining trust in institutions, polarization regarding pandemic measures, and concern about the broader effects of polarization, offer a comprehensive view of public attitudes and sentiments during a global crisis. They can provide valuable insights for policymakers, public health officials, and researchers aiming to navigate the complex socio-political landscape of the COVID-19 pandemic and similar future events.

Moreover, we collected following measures only at Wave 1 as we consider them as time invariant as they either measure on experience during the pandemic or stable traits, Personal affectedness (Hansen, S., Schäfer, N., & Kaspar, R. (2023)) and BIG 5 personality scale (Rammstedt and John 2007). Personal affectedness refers to the degree to which individuals perceive themselves to be personally impacted by an issue or situation. This concept encompasses emotional, cognitive, and behavioral dimensions, indicating how strongly a person feels directly involved or likely to experience the consequences of a given event or condition. It's a measure that captures the subjective intensity of personal involvement, which can significantly influence attitudes, decision-making processes, and behaviors in response to specific contexts or challenges. The Short BIG 5 personality scale is a condensed instrument designed to assess the five major dimensions of personality traits efficiently: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. We used this scale as time constraints were an important aspect during the assembly.

Variables collected only at Wave 3 evaluation_measures

5.4 A timeline of the experiment

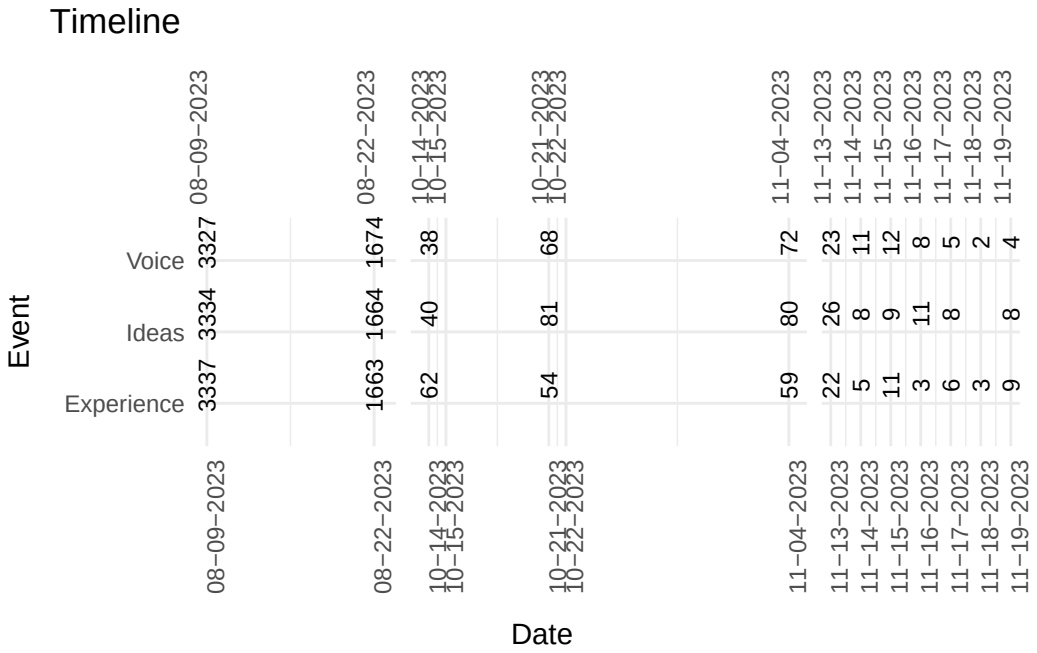

Figure 1: Timeline of the experiment

6 Results

We will structure our results in a Invitation part and an Assembly part. In the Invitation part we will present the results of the invitation experiment, while in the Assembly part we will present the results of the assembly experiment. Subjects who received an invitation we will refer to as invitees, those who accepted the invitation as applicants, and those who participated in the assembly as participants. We have $invitees \subseteq applicants \subseteq participants$.

6.1 Invitation stage

We first present descriptive statistics about the invitees. Table 1 shows the distribution of participant characteristics across the different treatments. The distribution on observables reflects the (marginal) distributions of these characteristics of the Austrian population in the age range of 18 to 89. Treatment allocation is balanced over those observables.

Table 1: Descriptive statistics about participants in a citizen assembly experiment

		Voice (N=5001)		Ideas (N=4998)		Experience (N=5000)	
		N	Pct.	N	Pct.	N	Pct.
Gender	female	2593	51.8	2589	51.8	2590	51.8
	male	2408	48.2	2409	48.2	2410	48.2
Age group	18-34	1145	22.9	1146	22.9	1146	22.9
	35-59	2075	41.5	2075	41.5	2075	41.5
	60-89	1781	35.6	1777	35.6	1779	35.6
Education	No Matura	3930	78.6	3930	78.6	3927	78.5
	Matura	1071	21.4	1068	21.4	1073	21.5

6.1.1 Treatment effects

We now turn to our main treatment effects on participation. Table 2 shows the acceptance of the invitation regressed on our treatments. We use a logistic regressions, weighting by the size of the population of each state and sequentially adding state fixed effects and covariates to assess robustness. We observe that in the more cognitive demanding treatments *Ideas* and *Experience* the acceptance rate is lower than in the *Voice* treatment. To assess the statistical significance of the differences.

We find a robust positive effect of our treatment **Voice** compared to the baseline **Ideas** treatment on the number of subjects who stated the willingness to participate. Moreover, acceptance rates in the **Voice** treatment is also significantly higher that in the **Experience** treatment, partially confirming our Conjecture 3.1. The magnitude of the effect of the **Voice** treatment is around 17~percent when controlling for all available covariates and adding state level fixed effects. The effect is robust to the inclusion of state fixed effects and covariates. Moreover, as expected, we see that invitees with a high-school degree have a higher acceptance rate than those without a high-school degree and men have a higher acceptance rate than, although the latter is not robust to the inclusion of state fixed effects and covariates.

6.1.2 Education

We will now go through subgroups as defined by our observables in the first stage. First, we saw a strong positive correlation between invitation acceptance and attending high school. In Table 3 we present the results of the same logistic regression of the acceptance of the invitation on the treatment interacted with education, to assess the differential effects of the treatments

Table 2: Logit model acceptance of invitation treatments

	(1)	(2)	(3)	(4)
Panel A: Main regression				
Treat: Voice	0.176***	0.181***	0.177*	0.181*
	(0.000)	(0.000)	(0.083)	(0.088)
Treat: Experience	0.023***	0.015***	0.023	0.014
	(0.000)	(0.000)	(0.103)	(0.104)
High school		0.879***		0.865***
		(0.000)		(0.087)
Male		0.106***		0.103
		(0.000)		(0.075)
Num.Obs.	14999	14999	14999	14999
FE: state			X	X
Panel B: Contrasts				
Treat: Voice=Experience	0.154***	0.165***	0.154+	0.167+
	(0.000)	(0.000)	(0.089)	(0.089)
Num.Obs.	14999	14999	14999	14999

Panel A reports the main regression and Panel B the comparisons between treatment Experience and Voice. We applied state population weights and state fixed effects. Significance is reported at the following levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3: Logit model acceptance of invitation treatments interacted with education

	sample: Full sample	sample: 18-34	sample: 35-59	sample: 60-89
Treat: Voice x No High school	-0.131+	0.184	-0.129	-0.341**
	(0.068)	(0.266)	(0.174)	(0.110)
Treat: Voice x High school	0.851***	0.728***	0.772***	1.195***
	(0.087)	(0.193)	(0.174)	(0.225)
Treat: Experience x No High school	-0.210*	0.145	-0.193	-0.483+
	(0.107)	(0.156)	(0.133)	(0.255)
Treat: Experience x High school	0.548***	0.597*	0.609***	0.352
	(0.113)	(0.234)	(0.157)	(0.381)
Male	0.094	-0.128	-0.049	0.471***
	(0.075)	(0.259)	(0.105)	(0.137)
Age: 35-59	0.051			
	(0.061)			
Age: 60-89	-0.086			
	(0.101)			
Num.Obs.	14999	3437	6225	5337
FE: state	X	X	X	X

We applied state population weights and state fixed effects. Significance is reported at the following levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

on the acceptance of the invitation by education level. Contrary to our conjecture, we observe a negative effect of both the *Voice* and *Experience* treatments on the acceptance of the invitation for the less educated population. This seem to be driven mainly by subjects in the age group of 60-89.

The *Voice* and *Experience* treatments were particularly apt to attract applicants for subjects with high-school degree. The treatment *Voice* was able to attract more subjects over all age groups, especially the sample of the older subjects between 60-89.

6.1.3 Gender

There is a well documented literature on the differences in the willingness to participate in political activities by gender (e.g., M. Gerber, Schaub, and Mueller (2019) for an example of a Swiss assembly). Table 4 shows that the treatment *Voice* was able to attract more subjects over all gender and age groups, especially women in the age range of 18 to 34, while we a

Table 4: Logit model acceptance of invitation treatments interacted with gender

	Full sample	Age: 18-34	Age: 35-39	Age: 60-89
Treat: Voice x Female	0.159	0.551*	0.230	-0.227
	(0.105)	(0.218)	(0.181)	(0.197)
Treat: Voice x Male	0.207**	0.124	0.146	0.352**
	(0.079)	(0.205)	(0.114)	(0.112)
Treat: Experience x Female	-0.082	0.183	0.076	-0.520*
	(0.098)	(0.297)	(0.192)	(0.258)
Treat: Experience x Male	0.113	0.370	0.090	-0.093
	(0.153)	(0.256)	(0.089)	(0.267)
High school	0.852***	0.530**	0.757***	1.291***
	(0.079)	(0.181)	(0.162)	(0.121)
Age: 35-59	0.044			
	(0.064)			
Age: 60-89	-0.059			
	(0.098)			
Num.Obs.	14999	3437	6225	5337
FE: state	X	X	X	X

We applied state population weights and state fixed effects. Significance is reported at the following levels: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

non-significant negative point estimate for the treatment *Voice* women in the age group of 60-89 who also have a negative point estimate and significant for the treatment *Experience*.

We find that the treatment "Voice"was especially able to attract women in the age range of 18-34 and men in the age range of 60-89, compared to the baseline Ideas treatment. The treatment experience did attract less women in the sample of the subjects between 60-89.

6.2 Assembly data

We now turn to the results from the data collected at the assembly stage. We will first present the results from the survey conducted at the assembly, and then the results from the survey conducted after the assembly, in relation to our invitation intervention.

We are interested here whether and how our intervention has influenced the application of subjects concerning their locus of control. As this is argued to be a stable trait (i.e., not easily

Table 5: Fixed effects OLS model of locus of control depending on treatment

	(1)	(2)	(3)	(4)
Panel A: Main regression				
Treat: Experience	-0.103	-0.114	-0.118	-0.128+
	(0.080)	(0.067)	(0.080)	(0.069)
Treat: Voice	-0.262*	-0.129*	-0.289**	-0.181*
	(0.082)	(0.048)	(0.085)	(0.066)
Num.Obs.	313	316	312	315
FE: state	X	X	X	X
FE: gender			X	X
FE: education			X	X
Panel B: Contrasts				
Treat: Voice=Experience	0.159+	0.015	0.171+	0.053
	(0.094)	(0.071)	(0.096)	(0.077)
Num.Obs.	313	316	312	315

Panel A reports the main effects of the treatment on the locus of control, Panel B shows the effect differences between the treatments Experience and Voice. We applied state population weights and state fixed effects. Significance is reported at the following levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

influenceable by simple environmental cues Cobb-Clark and Schurer (2013)) we consider that it has not been shifted but that different types of people have applied. Table 5 shows that participants who were recruited through the voice treatment have a lower locus of control-–external and overall-–compared to the baseline treatment. This result is robust to controlling for state, age and education.

The differences between the baseline treatment and the *Experience* treatment are negative, but not robustly significant, so are the differences between the *Experience* and the *Voice* treatments.

We now turn to the comparison of our locus of control measured in the reference survey and in our assembly. Table 6 shows that the locus of control is higher in the citizen assembly experiment compared to the reference survey. This is consistent with our conjecture that citizens assemblies usually attract subjects who have a higher control belief and it is consistent with the general observation that subjects who participate in assemblies vary in the expression of their personality traits. Jennstål (2018) e.g. found that extraversion and openness are positively associated with the willingness to participate. Interestingly we do not find this in

Table 6: Descriptive statistics of locus of control in the reference survey and the citizen assembly experiment.

	Assembly (N=316)		Survey (N=2603)			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	Std. Error
Locus control overall	3.95	0.64	3.50	0.71	−0.45***	0.04
Locus control external	3.54	0.90	3.12	0.98	−0.42***	0.06
Locus control internal	4.34	0.72	3.88	0.82	−0.46***	0.04

our data, as shown in Table 10.

As **?@fig-loc-cs-cdf** shows, this difference in means is not driven by some extreme values but is persistent over the whole distribution. Our treatments first order stochastically dominate the distribution of LoC from the general population. A Kolmogorov-Smirnoff test of the survey measures of LoC and the experiment measures (pooled over all treatments as well as for each treatment separately) of LoC confirms that the two distributions are different (p-values: $<$ .001). Also a Hansen-Bowers test confirms that the distributions are different and we do not achieve balance over LoC, not even with the *Voice* treatment (p-values: $<$ .001). The effect size of the difference between the control survey and the treatments is 0.260 for the external LoC and 0.119 for the overall LoC for the voice treatment compared with the experience treatment this constitutes a small effect, however, given the very minimal zero-cost intervention, it can be considered important and stronger interventions in this directions may be futile.

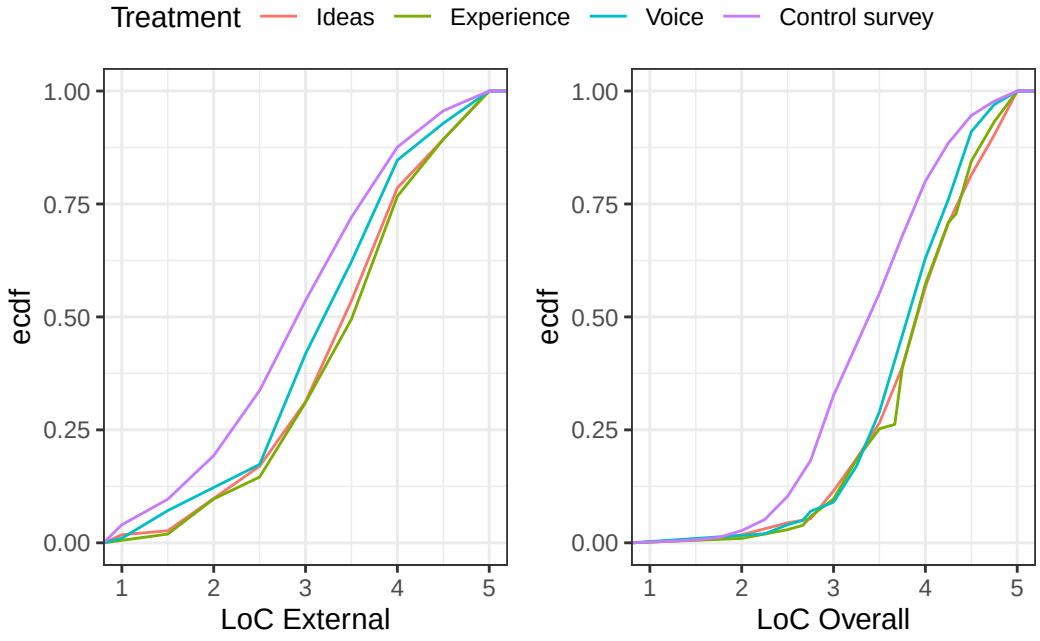

Table 7: Linear regression locus of control on treatments.

	LoC: External	LoC: External	LoC: Overall	LoC: Overall
Treat: Experience	-0.103	-0.099	-0.114	-0.099
	(0.080)	(0.085)	(0.067)	(0.079)
Treat: Voice	-0.262*	-0.255**	-0.129*	-0.137*
	(0.082)	(0.068)	(0.048)	(0.056)
Male		0.062		-0.001
		(0.150)		(0.080)
Diverse		-1.437***		-0.855***
		(0.120)		(0.084)
Age: 35-59		0.110		0.120
		(0.088)		(0.103)
Age: 60-89		0.222		0.255+
		(0.159)		(0.132)
genderNaN				1.147***
				(0.082)
Num.Obs.	313	312	316	315
R2	0.038	0.061	0.026	0.073

We applied state population weights and state fixed effects. Significance is reported at the following levels: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

We now look at the robustness of the effects. Table 7 confirms that the result are robust when controlling for age and gender.

When turning to the (non-preregistered) heterogeneous effects presented in Table 8 we observe an interesting pattern. For the Voice treatment we have pretty consistent negative effects on LoC in all age groups and over gender, with one exception, females in the 60-89 age group.

For the Experience treatment we have a negative effect on LoC for the 35-59 age group and a positive effect for the 60-89 age group that seems to be mainly driven by male participants, that also have a higher LoC in the age group of 18-34.

6.2.1 Distributions of personal characteristics by treatment

We now will turn to the distribution of further personal characteristics by treatment. We start by looking at whether the subject has been individually affected by the pandemic on the dimen-

Table 8: Linear regression locus of control on treatments by gender and age group.

	All data		Female		
	LoC: External	LoC: Overall	LoC: External	LoC: Overall	LoC: Exter
Treat: Experience x Age: 18-34	-0.088	-0.142	0.517+	0.284	-0.626*
	(0.178)	(0.186)	(0.237)	(0.173)	(0.201)
Treat: Experience x Age: 35-59	-0.264*	-0.190*	-0.281	-0.164	-0.242
	(0.095)	(0.078)	(0.172)	(0.178)	(0.191)
Treat: Experience x Age: 60-89	0.253*	0.058	0.243	0.030	-0.060
	(0.082)	(0.149)	(0.249)	(0.165)	(0.194)
Treat: Voice x Age: 18-34	-0.438*	-0.309*	-0.300	-0.282	-0.801**
	(0.149)	(0.100)	(0.239)	(0.174)	(0.135)
Treat: Voice x Age: 35-59	-0.267	-0.170	0.134	0.144	-0.829**
	(0.168)	(0.100)	(0.249)	(0.160)	(0.233)
Treat: Voice x Age: 60-89	-0.123	0.002	0.218	0.241+	-0.310
	(0.209)	(0.112)	(0.192)	(0.126)	(0.295)
Num.Obs.	311	313	159	160	152
R2	0.056	0.046	0.083	0.073	0.215

We applied state population weights and state fixed effects. Significance is reported at the following levels: + ** $p < 0.01$, *** $p < 0.001$.

Table 9: Linear regression of being personally affected on treatments.

	Physically	Mentally	Financially	Job	Hobbies	Socially	Freedom
Treat: Experience	0.354*	0.286*	-0.055	0.298*	-0.107	-0.172	0.092
	(0.120)	(0.120)	(0.174)	(0.094)	(0.232)	(0.232)	(0.282)
Treat: Voice	0.219	0.091	0.331	0.085	0.016	0.035	0.119
	(0.235)	(0.085)	(0.298)	(0.226)	(0.247)	(0.172)	(0.198)
Num.Obs.	315	315	311	311	313	315	316
R2	0.030	0.035	0.032	0.021	0.022	0.011	0.031

We applied state population weights and state fixed effects. Significance is reported at the following levels: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

Table 10: Linear regression of BIG 5 personality traits on treatments.

	Extrov.	Agree.	Consci.	Neuro.	Open.
Treat: Experience	0.031	0.040	0.026	0.056	-0.048
	(0.246)	(0.096)	(0.059)	(0.069)	(0.090)
Treat: Voice	0.093	-0.182	0.097	0.169	0.031
	(0.252)	(0.118)	(0.060)	(0.137)	(0.089)
Num.Obs.	315	314	314	311	315
R2	0.034	0.034	0.028	0.024	0.017

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

sions: Physically, mentally, financially, within the job, within free time and hobbies, socially, and personal freedom. Table 9 shows that the *Experience* treatment has a significant effect on attracting the subject who have been physically or mentally affected more severely. The point estimates of the *Voice* treatment point in a similar direction, but are not significant.

Finally we will have a look at the BIG 5 personality traits as Jennstål (2018). Table 10 again shows no significant differences in the personality structure of the subjects attracted by the different treatments. The point estimates are small and not significant.

7 Conclusion

We conducted an experiment to study the effect of three different conditions on the willingness of Austrian citizens to participate in a citizen assembly and to assess the the differences in obserables and unobservables. We find that language that empahsizes the inclusive nature of

the event attracts 17% more subjects. Moreover theses subjects are more likely to have a lower cost of control on average.

We can have an optimal mix of the three conditions by stratifying over age and gender. This will allow us to have a balanced sample of the three conditions with the knowledge.

Table A1

	sample: Full sample	sample: BUR	sample: KAE	sample: NIE	sample: OBE	sam
Treat: Voice	0.070	−0.048	0.023	0.142	0.066	
	(0.084)	(0.293)	(0.267)	(0.280)	(0.285)	
Treat: Experience	0.004	0.003	−0.235	−0.336	−0.003	
	(0.094)	(0.290)	(0.282)	(0.316)	(0.291)	
Male	0.065	0.237	−0.308	0.225	−0.278	
	(0.072)	(0.239)	(0.228)	(0.242)	(0.238)	
Age: 35-59	0.009	−0.123	−0.213	−0.114	0.205	
	(0.065)	(0.316)	(0.276)	(0.293)	(0.295)	
Age: 60-89	−0.178+	−0.259	−0.531+	−0.611+	−0.196	
	(0.098)	(0.324)	(0.293)	(0.337)	(0.330)	
Num.Obs.	14 999	1791	1631	1791	1631	
FE: state	X	X	X	X	X	

Significance is reported at the following levels: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

A Appendix

A.1 Results by State

In Table A1 we present the results of the logistic regression of the acceptance of the invitation on the treatment by state.

B Text Data merging procedure

Data from time point 1 and 2 were merged on November 6th, 2023, after the last workshop. First, the submissions from the survey platform Limesurvey were merged. On time point 1 (2), 154 (165) full responses were submitted, thereof 11 (8) incomplete responses and 2 (4) duplicated responses were deleted. In case of duplicated responses, the second entry was deleted. Using R-studio, 130 participants were automatically matched by their identification code, 34 remained unmatched. Second, paper-pencil responses were typed into excel by 2 research assistants, and again using R-studio merged by their identification code state by state. 149 entries were automatically matched, 44 remained unmatched. Thus, overall, 279 participants were automatically merged. Then, 35 participant entries were merged manually based on their identification code, age, and gender, as these participants used different formats

in time point 1 and 2 (i.e., paper-pencil and online format). In sum, this leaves us with 322 observations of which 314 are matched in time point 1 and 2 and 8 unmatched observations. A majority were paper-pencil responses (N = 157) or online responses (N = 130) and some mixed (N = 35). Assessment of merge. We assessed correct matching by checking whether age and gender, which were collected in both time points, are identical. Moreover, matching the paper-pencil responses state by state allowed to check for the state and avoid wrong matches. We excluded participants who did not complete the survey on the days of data collection, except for 1 case, who finished the second survey one day after workshop. Participants' comments were screened for unsuitable responses (e.g., this is just a test, I am not the person invited). Data correction. In 35 cases, the identifier codes contained typos. These were corrected manually based on gender, age, and state to allow automatic merging. In the pencil-paper questionnaires, the research assistants report 73 unclear entries when typing into excel. Thereof, 27 could be unambiguously corrected by the authors (36.99%), the remaining were set to missing values. In 24 cases, the age entry was corrected by the authors, as participants entered either no age at one time point or a clear typo was obvious (e.g., 85 instead of 58 years). Same procedure was applied to 7 entries for gender. In the paper-pencil questionnaire, job status was set to missing values if more than one response was selected. Control group. In the control group, we had 81 online responses, of which 68 were complete. Thereof, 2 test responses and 2 which did not submit in the given timeframe were excluded. Thus, 62 participants were in the control group in time point 1. Opposed to the pre-registration, we did not exclude participants who filled out the survey during weekdays.

Data from *time point 3*, were merged in a similar procedure after the time frame for filling out the questionnaire online or via pencil paper. Paper and online were merged by intervention and control separately. Of the 269 initial submissions in the intervention group (thereof 18 pencil paper) and the 116 submissions in the control group (thereof 19 pencil paper), we excluded 37 duplicates responses due to repeated participation in the intervention group and two in the control group, while always keeping the first record. Many of those participants again filled the survey when receiving the reminders. Also, we excluded one case, where the pencil paper questionnaire arrived on 29.11.2023 and two test responses. Thus, we have 345 observations at time point three, of which 242 are in the intervention group and 103 are in the control group.

Data correction. As the final reminders were send incorrectly (intervention group got reminder for control group in 6 states), we had 22 cases in the control group, where affiliation to treatment group/control group was unclear. Based on overall merge of all time points in the end, we were able to correctly assign the respective group, based on code, age, gender, and state. We made changes in reported age, gender in 20 cases. Most of those cases were changes in age, as several individuals reported an age increase of one year at time point 3 (birthdays). In the pencil-paper questionnaires, the research assistants report 5 unclear entries when typing into excel, which were set to missing values.

Overall merge. We again merged the data from time point 1 and 2 with the data from time point 3 based on code, and corrected successful merging based on gender, age, and state. The

remaining cases were only merged at two timepoints or not merged at all or in the control group. In total, we merged 4 cases manually und unmerged two cases which were merged incorrectly based on the code. Finally, we changed 5 codes to avoid duplicate codes.

In sum, we have 488 complete and incomplete observations, of which 368 are in the intervention group and 120 are in the control group. Of the 368 observations in the intervention group, 192 observations in the intervention group were merged across all three time points. Of the 120 observations in the control group, 45 were merged across time points 1 and 3. 251 observations remain without complete match. We created one variable each for gender, age, state, and control group for all three times points.

C Participant Survey 1

This survey was administered at the day of the assembly before the start of deliberation.

C.1 "Österreich am Wort" Dialogue Process Questionnaire

Dear participant!

The "Österreich am Wort" dialogue process is being scientifically monitored. There are three questionnaires: one before the event (#1), one at the end (#2), and one approximately 2 weeks after the "Österreich am Wort" event (#3).

In the first questionnaire, we are interested in your attitudes towards various current topics. It takes about 8 minutes to complete.

Please answer spontaneously and honestly. There are no right or wrong answers. Only your opinion counts! Please complete the questionnaire as fully and accurately as possible. You will be making a significant contribution to research and the management of future crises.

Thank you very much for your support!
Katharina Gangl
Contact: gangl@ihs.ac.at

Group Number:
Please enter the group number you received when you registered:

Individual Code:
Please turn the page, the questionnaire starts on the back! As we will be conducting the survey several times, we ask you to create an individual code each time. This allows us to link the data while maintaining your anonymity. We cannot deduce your identity from the code.

- What is the first letter of your mother's first name? (Example: Ursula - U):
- What is the first letter of your father's first name? (Example: Ahmet - A):
- What month is your birthday? (Example: June = 06):
- What is the first letter of your current place of residence? (Example: Bad Gastein - B):
- How many older siblings do you have?

Trust Evaluation:

Please indicate how strongly you agree with the following statements:

I have trust in…	Completely disagree	Rather disagree	Neither/nor	Rather agree	Completely agree
… public service media (TV, newspapers)					
… private media (TV, newspapers)					
… social media					
… local media					
… public research institutions (universities)					
… private research institutions (companies)					
… the European Union (EU)					
… the Austrian federal government					
… the state government					
… the politics of my hometown					
… people in Austria					
… people in my state					
… people in my neighborhood					

Evaluation of People's Experience with Corona Measures:

Think about the corona measures. How do you currently experience people who supported these measures at the time:

For each term, please indicate how much you think it applies to these people:

Term	Completely disagree	Rather disagree	Neither/nor	Rather agree	Completely agree
Responsible					
Intelligent					
Honest					
Generous					
Trustworthy					

Think about the corona measures. How do you currently experience people who opposed these measures at the time:

For each term, please indicate how much you think it applies to these people:

Term	Completely disagree	Rather disagree	Neither/nor	Rather agree	Completely agree
Responsible					
Intelligent					
Honest					
Generous					
Trustworthy					

Impact of the Corona Period:

Some say that the corona period has politically divided society and led to more polarization. Others say that the corona period has united society and reduced polarization. And yet others say that the corona period has had no impact. Please select one of the following answers:

The corona period…

has significantly reduced the polarization of society
has slightly reduced the polarization of society
had no impact
has slightly increased the polarization of society
has significantly increased the polarization of society

How much do you agree with the following statement:
„I am concerned about the polarization between opponents and supporters of the coronavirus measures.“

Completely disagree	Rather disagree	Neither/nor	Rather agree	Completely agree

---

Personal Preferences and Traits:

Now we are interested in you and your preferences. The following statements may apply to you to a greater or lesser extent. For each statement, please indicate the extent to which it applies to you:

Statement	Does not apply at all	Does not apply much	Neither/nor	Applies somewhat	Applies fully and completely
If I work hard, I will succeed.					
I'm my own boss.					
Whether at work or in my private life: What I do is mainly determined by others.					
Fate often gets in the way of my plans.					

How much did you feel negatively affected by the coronavirus pandemic in the following areas:

Area	Not affected at all	Rather not affected	Neither/nor	Rather affected	Very affected
Physical Health					
Mental Health (Mood/Feelings/Well-Being)					
Financial Situation					
Career					
Leisure Activities/Hobbies					
Social contacts and Personal Relationships					
Personal Freedom					

To what extent do the following statements apply to you:

Statement	Does not apply at all	Does not apply much	Neither/nor	Applies somewhat	Applies fully and completely
… is reserved					
… is generally trusting					
… tends to be lazy					
… is relaxed, handles stress well.					
… has few artistic interests					
… is outgoing, sociable					
… tends to find fault with others					
… does a thorough job					
… gets nervous easily					
… has an active imagination					

Personal Information:

- Which gender would you categorize yourself as? Female
  Male
  Diverse
- How old are you:
- Please state the highest level of education you have completed: No compulsory school leaving certificate
  Matura
  Compulsory school
  Master Craftsman or Foreman
  Apprenticeship with vocational school
  University or other Higher Education
  Specialized- or Trade School
  Other degree after Matura
- What is your current professional status: Employed
  Work in the household
  Self-employed
  Retired
  Job seeking

Others
In training (study/school)

- How many inhabitants does your residential municipality have: Up to 5,000 inhabitants
  5,001 to 50,000 inhabitants
  50,001 to 100,000 inhabitants
  More than 100,000 inhabitants

Is there anything else you would like to tell us at the end?

Thank you very much for your participation! Your contribution helps us a lot.

References

Al-Ubaydli, Omar, John A. List, Danielle LoRe, and Dana Suskind. 2017. "Scaling for Economists: Lessons from the Non-Adherence Problem in the Medical Literature." *Journal of Economic Perspectives* 31 (4): 125–44. https://doi.org/10.1257/jep.31.4.125.

Andor, Mark A., James Cox, Andreas Gerster, Michael Price, Stephan Sommer, and Lukas Tomberg. n.d. "Locus of Control and Prosocial Behavior." Working Paper. https://doi.org/10.3386/w30359.

Andreoni, James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *Economic Journal* 100 (401): 464. https://doi.org/10.2307/2234133.

Bächtiger, André, John S. Dryzek, Jane Mansbridge, Mark E. Warren, André Bächtiger, John S. Dryzek, Jane Mansbridge, and Mark E. Warren, eds. 2018. *The Oxford Handbook of Deliberative Democracy*. Oxford Handbooks. Oxford, New York: Oxford University Press.

Ban, Radu, Saumitra Jha, and Vijayendra Rao. 2012. "Who Has Voice in a Deliberative Democracy? Evidence from Transcripts of Village Parliaments in South India." *Journal of Development Economics* 99 (2): 428–38. https://doi.org/10.1016/j.jdeveco.2012.05.005.

Boxell, Levi, Jacob Conway, James N. Druckman, and Matthew Gentzkow. 2022. "Affective Polarization Did Not Increase During the COVID-19 Pandemic." *Quarterly Journal of Political Science* 17 (4): 491–512. https://doi.org/10.1561/100.00021027.

Carroll, Rory. 2024. "Irish Referendum Fiasco Puts Future of Lauded Citizens' Assemblies in Doubt." *The Guardian*, March 20, 2024, sec. World news. https://www.theguardian.com/world/2024/mar/20/irish-referendum-fiasco-puts-future-of-lauded-citizens-assemblies-in-doubt.

Chattopadhyay, Raghabendra, and Esther Duflo. 2004. "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India." *Econometrica* 72 (5): 1409–43. https://www.jstor.org/stable/3598894.

Cobb-Clark, Deborah A., and Stefanie Schurer. 2013. "Two Economists' Musings on the Stability of Locus of Control." *The Economic Journal* 123 (570): F358–400. https://doi.org/10.1111/ecoj.12069.

Cox, D. R. 1958. *Planning of Experiments*. Planning of Experiments. Oxford, England: Wiley.

Doherty, Carroll, Jocelyn Kiley, and Nida Asheer. 2019. "Partisan Antipathy: More Intense, More Personal." *Pew Research Center*.

Dryzek, John S., and Simon Niemeyer. 2008. "Discursive Representation." *American Political Science Review* 102 (4): 481–93. https://doi.org/10.1017/S0003055408080325.

Fishkin, James, Valentin Bolotnyy, Joshua Lerner, Alice Siu, and Norman Bradburn. 2024. "Can Deliberation Have Lasting Effects?" *American Political Science Review*, February, 1–21. https://doi.org/10.1017/S0003055423001363.

Gatz, Margaret, and Michele J. Karel. 1993. "Individual Change in Perceived Control over 20 Years." *International Journal of Behavioral Development* 16 (2): 305–22. https://doi.org/10.1177/016502549301600211.

Gerber, Alan S., Gregory A. Huber, David Doherty, Conor M. Dowling, Connor Raso, and Shang E. Ha. 2011. "Personality Traits and Participation in Political Processes." *Journal of Politics* 73 (3): 692–706. https://doi.org/10.1017/S0022381611000399.

Gerber, Marlène, Hans-Peter Schaub, and Sean Mueller. 2019. "O Sister, Where Art Thou? Theory and Evidence on Female Participation at Citizen Assemblies." *European Journal of Politics and Gender* 2 (2): 173–95. https://doi.org/10.1332/251510819X15471289106095.

Hibbing, John R., and Elizabeth Theiss-Morse. 2002. *Stealth Democracy: Americans' Beliefs About How Government Should Work*. Cambridge Studies in Public Opinion and Political Psychology. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511613722.

Jacquet, Vincent. 2017. "Explaining Non-Participation in Deliberative Mini-Publics." *European Journal of Political Research* 56 (3): 640–59. https://doi.org/10.1111/1475-6765.12195.

Jennstål, Julia. 2018. "Deliberative Participation and Personality: The Effect of Traits, Situations, and Motivation." *European Political Science Review* 10 (3): 417–40. https://doi.org/10.1017/S1755773918000024.

Kovaleva, Anna, Constanze Beierlein, Christoph J. Kemper, and Beatrice Rammstedt. 2012. "Eine Kurzskala Zur Messung von Kontrollüberzeugung: Die Skala Internale-Externale-Kontrollüberzeugung-4 (IE-4)." GESIS-Working papers 2012/17. GESIS.

Lacelle-Webster, Antonin, and Mark E. Warren. 2021. "Citizens' Assemblies and Democracy." In *Oxford Research Encyclopedia of Politics*. https://doi.org/10.1093/acrefore/9780190228637.013.1975.

Landemore, Helene. 2017. *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton, NJ Oxford: Princeton University Press.

Lohmann, Susanne. 1993. "A Signaling Model of Informative and Manipulative Political Action." *American Political Science Review* 87 (2): 319–33. https://doi.org/10.2307/2939043.

———. 1994. "Information Aggregation Through Costly Political Action." *The American Economic Review* 84 (3): 518–30. https://www.jstor.org/stable/2118065.

Maniadis, Zacharias, Fabio Tufano, and John A. List. 2014. "One Swallow Doesn't Make a Summer: New Evidence on Anchoring Effects." *American Economic Review* 104 (1): 277–90. https://doi.org/10.1257/aer.104.1.277.

Neblo, Michael A., Kevin M. Esterling, Ryan P. Kennedy, David M.J. Lazer, and Anand E.

Sokhey. 2010. "Who Wants to Deliberate—and Why?" *American Political Science Review* 104 (3): 566–83. https://doi.org/10.1017/S0003055410000298.

Osborne, Martin J, Jeffrey S Rosenthal, and Matthew A Turner. 2000. "Meetings with Costly Participation." *American Economic Review* 90 (4): 927–43. https://doi.org/10.1257/aer.90.4.927.

Paulhus, Delroy. 1983. "Sphere-Specific Measures of Perceived Control." *Journal of Personality and Social Psychology* 44 (6): 1253–65. https://doi.org/10.1037/0022-3514.44.6.1253.

Rammstedt, Beatrice, and Oliver P. John. 2007. "Measuring Personality in One Minute or Less: A 10-Item Short Version of the Big Five Inventory in English and German." *Journal of Research in Personality* 41 (1): 203–12. https://doi.org/10.1016/j.jrp.2006.02.001.

Riener, Gerhard, and Katharina Gangl. 2023. "Citizen Assembly Austria 2023." https://doi.org/10.17605/OSF.IO/RBFGQ.

Riener, Gerhard, Sebastian O Schneider, and Valentin Wagner. n.d. "Addressing Validity and Generalizability Concerns in Field Experiments."

Rotter, Julian B. 1966. "Generalized Expectancies for Internal Versus External Control of Reinforcement." *Psychological Monographs: General and Applied* 80 (1): 1–28. https://doi.org/10.1037/h0092976.

Vivalt, Eva. 2020. "How Much Can We Generalize From Impact Evaluations?" *Journal of the European Economic Association* 18 (6): 3045–89. https://doi.org/10.1093/jeea/jvaa019.