11
Experimental Methods

Kimmo Grönlund and Kaisa Herne

Early literature on deliberative democracy is largely normative. That is, it describes an ideal model of democracy. However, more recent research has also engaged in empirically testing the possibilities of achieving deliberation in actual political discourse and decision-making. One of the empirical methods in the deliberative democracy literature is experimentation. While experimentation in political science is by no means a novel invention (Gosnell 1926), it has seen a rapid rise over the last few decades (Morton and Williams 2010). At first sight, it might seem problematic to test something that is considered to present a normative model. However, experimentation on deliberative democracy aims at *testing the consequences of institutions that include deliberation*, and thereby providing a contribution to the *design of democratic institutions*. More broadly, like any normative theory, deliberative democracy has empirical elements, it makes certain claims about what would happen if ideal-type discussion were to take place, and such claims can be tested.

Experiments in deliberative democracy have mainly concerned deliberative *mini-publics*, that is, more or less representative samples of the public deliberating on specific political issues. When we talk about deliberation or deliberative discussion, we refer to organized discussion that meets at least some minimum requirements, such as facilitated discussion that takes place in small groups, and that is guided by discursive rules that emphasize respect for others, listening to others, and justification of one's views (e.g. Grönlund et al. 2015). The main questions addressed in experiments have been whether and to what extent deliberative mini-publics achieve the kinds of outcomes that deliberative discussion is supposed to produce. Are opinions changed? Are they changed into a more tolerant direction? Is trust increased? Is the legitimacy and acceptance of decisions increased? In other words, would introducing the deliberative element into democratic institutions make democracies stronger?

The chapter proceeds as follows. We first introduce the main characteristics of an experiment. Thereafter, we present the main object of experimental study in deliberative democracy, that is, deliberative mini-publics. That section takes up the strengths and challenges of engaging lay citizens in experiments, and specifically addresses data generation and analytical methods related to citizen deliberation. Finally, we

Kimmo Grönlund and Kaisa Herne, *Experimental Methods.* In: *Research Methods in Deliberative Democracy.* Edited by Selen A. Ercan, Hans Asenbaum, Nicole Curato, and Ricardo F. Mendonça, Oxford University Press.
© Oxford University Press (2022). DOI: 10.1093/oso/9780192848925.003.0011

discuss the interpretations that can be made in relation to the theory of deliberative democracy based on experiments, and address some limitations of the experimental method.

What Is an Experiment?

The reason for using experiments is that they provide a control of variables not available through naturally occurring data, and that this control makes it possible to detect causal mechanisms. For example, if we want to know how participation in a mini-public influences individual opinions, then organizing an experiment where some of the subjects participate in a mini-public, whereas others do not, allows for a comparison between these two groups. By this design, we are able to measure whether and how taking part in a mini-public influences opinions. Of course, real-world cases in which some people participate and others do not can also be observed. However, only in an experiment can we hold everything else between the two groups—those who deliberate and those who do not—constant, and thereby be confident that observed differences between the groups are due to experimental manipulation and not some other differences between the groups.

An experiment is a controlled study of a *causal relationship*. By experimentation we can examine the influence of an independent variable on one or several dependent variables. In experiments, independent variables are those variables that are manipulated. They could also be called predictor variables. Dependent variables, in turn, are not manipulated. They are outcome variables whose values depend on the manipulated, that is, independent, variables. To take an example, a mini-public experiment, Setälä et al. (2010) manipulated the small group decision-making method by using either a secret ballot or a consensual statement. In this study, the decision-making method was the independent variable and opinion formation the dependent variable. What distinguishes an experiment from other types of research that looks into causal relationships is the *manipulation* or *treatment* of an independent variable. An ideal experiment holds n-1 independent variables constant and manipulates only one. When other variables are held constant, we are able to say that any variation in the dependent variable is due to variation in the manipulated independent variable. Controlled variation is the key for detecting cause–effect relations.

For example, if we are interested in the influence of deliberative discussion on opinion change, we give a part of our subjects the *treatment*, that is, have them discuss with each other in conditions that approximate those of ideal deliberative discussion. Another part of the subjects, the *control group*, does not discuss under ideal deliberative conditions. We then measure participants' opinions *pre* and *post* taking part in deliberative discussion (the treatment group) or engaging in some other activity but not deliberating (the control group).

In the case of deliberation, the control group can take several forms: it can simply fill in the *pre* and *post* surveys with a certain time lag between the two surveys

(see Grönlund et al. 2015), or it can engage in free discussion without the elements of deliberative discussion (cf. Strandberg et al. 2019). The control group might also engage in some other activity, for example, writing down their ideas about the discussed topic individually. Since the treatment group is compared to the control group, the point of comparison varies depending on what the control group does. For example, if the control group engages in free discussion and the treatment group engages in deliberative discussion, we can detect the difference between these two types of discussions—what is the added value of discussion being deliberative? On the other hand, if the control group writes down their thoughts individually while the treatment group engages in deliberative discussion, we can detect the difference between deliberative discussion and individual reflection. Sometimes experimental designs lack the possibility of a control group and different treatments serve as one another's controls. For example, in the aforementioned study on the influence of the decision-making method on participants' opinion formation in a mini-public, the two treatment conditions were considered one another's controls (Setälä et al. 2010).

In addition to using some kind of treatment, a random assignment to treatment and control groups is crucial to experiments. Randomization is needed because it rules out systematic variation between treatment and control group participants. Without randomization we cannot be sure whether observed differences are due to the experimental treatment or to systematic variation between the participants in the treatment and control groups. It is noteworthy that participants in face-to-face mini-public experiments have to show up at a certain place and invest their time, as well as other resources, in deliberative discussion. It is obvious that this investment can influence the type of people who are likely to attend. For this reason, we cannot rule out systematic variation if randomization to treatment and control groups is done amongst all those invited to take part. A better strategy is to randomize amongst those who are *willing* to participate, or better still, if possible, amongst those who *show up*. The reason is that those who are invited, those who are willing to participate, and those who show up might represent slightly different types of groups. This means that we cannot be sure whether the observed outcome variation occurs because of our treatment or rather because of the differences between the types of people (not) willing to participate.

In addition to the control of variables, a clear advantage of experiments is that they can be *replicated*. Replication of research is not easily available through naturally occurring data. However, replication is crucial because it increases the reliability of observed results. For example, if we obtain similar results from mini-public experiments organized in different countries, we can be more confident about the generalizability of the results.

In social sciences, three different types of experiments can be separated. These types can roughly be characterized according to a scale from more control to less control, and from less external validity to more external validity. Laboratory experiments allow for most control but are often criticized for a lack of external validity. The critique usually stems from the use of student subjects and conditions that lack

similarity with real-world circumstances. For this reason, there has been a call for more realistic field experiments conducted in participants' natural environments. While field experiments bring about certain advantages, the drawback is that they cannot achieve the same degree of control as laboratory experiments. Lab-in-the-field experiments fall between pure field and pure laboratory experiments and are based on an idea of increasing external validity, for example by bringing experiments into natural locations without losing much of the control of variables. Mini-public experiments based on a random sample of the population and taking place either online or face to face commonly fall into the category of lab-in-the-field experiments. However, drawing the line between different types of experiments is not unambiguous.

While randomization is commonly seen as essential to experimentation, there are experiments that lack random assignment. In *quasi experiments*, subjects are not randomized into treatment and control groups. There are two types of quasi experiments, depending on whether the experimenter has control over the assignment of subjects into treatment and control conditions. In naturally occurring experiments, the experimenter has no control over the assignment of experimental subjects into treatment and control groups, but these groups are formed in a natural process. An example is a study which compares Brazilian municipalities that use or do not use participatory budgeting (Touchton and Wampler 2014). Naturally occurring experiments have the lowest level of control over variables, but they provide real-world cases that resemble experiments. When the experimenter has control over the assignment in a quasi-experimental design, subjects are allocated into treatment and control conditions based on a certain characteristic.

Experiments and Deliberative Democracy

Experimental research is usually motivated by a willingness to test theories. Based on the theory or theories, specific hypotheses are formulated. These are then tested with an appropriate experimental design. Theories are challenged or given support depending on the results of the experiment. However, in the case of deliberative democracy, the relationship between theory and empirical research is rather atypical because deliberative theory is a normative characterization of ideal states of affairs. While normative theories cannot be tested as such, they involve empirical claims that can be tested, and this is what experiments on deliberative democracy seek to achieve (Setälä and Herne 2014).

It is also notable that experiments are closely related to the central concepts of deliberative democratic theorizing. Experiments in deliberative democracy study the consequences of taking part in *deliberative discourse*. They often study how *opinions change* during deliberation, for example whether they become more similar or whether a meta-level consensus is achieved (Dryzek and Niemeyer 2006). Many

experiments also measure different *civic virtues* central to deliberative democracy, for example social and political trust, political efficacy, political knowledge, and readiness to participate in politics. Moreover, many experiments touch upon the *legitimacy* of decisions and procedures by asking participants how they conceive the decision-making processes and decisions made in them.

The usual motivation for experimentation is to test a given theoretical model. However, in the case of deliberative democracy, the motivation for experimentation can also be a study of discursive democratic innovations not yet used in real-world democratic systems. Via experiments, we can study the potential consequences of certain ways of organizing democratic participation, and, on the basis of the results, decide whether these ways should be put to use to improve the quality of democracy. Experiments can thereby contribute to the design of democratic institutions.

The focus of this chapter is on experiments with *deliberative mini-publics*. The reason for this is simple: experimental research on deliberative democracy has so far mainly concerned mini-publics. The reason for the popularity of mini-public experiments is likely their usefulness to test issues, such as opinion change, discussion dynamics, as well as various 'side effects' of deliberation, for example, political and social trust, political knowledge, and political efficacy (Grönlund et al. 2010; Grönlund 2016).

One factor that varies between different mini-publics is size, that is, the number of participants. Even though all mini-publics have the goal of being representative in the sense of representing different viewpoints amongst the public, most mini-publics fall short of being representative in a statistical sense and none of them are representative in the electoral sense (Goodin and Dryzek 2006). The best-known models for organizing mini-publics are Citizens' Juries (12–26 participants), Consensus Conferences (10–50 participants), Citizen Assemblies (50–160 participants), and Deliberative Polling (100–500 participants) (Breckon et al. 2019). Of these models, the most widely implemented mini-public is the Deliberative Poll™, developed by James Fishkin (1991). Deliberative Polls have by now been organized over 100 times in twenty-eight countries (Fishkin 2018). The Deliberative Poll model has become so widely known that Jane Mansbridge (2010) has dubbed it the 'gold standard' of how to organize deliberative mini-publics.

In many cases, Deliberative Polls and other mini-publics involve policymakers (Grönlund et al. 2014, 1). Involving external actors, such as national, regional, or local governments, is good if your aim is to make deliberative democracy 'bite', that is, to connect it with the democratic system. Yet, from a scientific viewpoint, it is likely to decrease the possibilities to design experimental manipulations freely. Since the scientific aim of most deliberative mini-publics is to study how the participants are affected by deliberation, this may decrease the internal validity of the mini-public as an experiment. Thus, for a researcher there is a trade-off between choosing to stimulate public debate (or even to have a real impact on democratic decision-making) on the one hand, and designing a coherent controlled experiment on the other. This

is something that the researcher should decide early in the planning process. By the time external actors are invited, it might be too late. Sometimes external actors are necessary for the simple reason that the researcher needs funding for her experiment.

In mini-publics, deliberation normally takes place in *small groups*; it is guided by *discussion rules*, and facilitated by *a trained moderator* (e.g. Grönlund et al. 2015). The rules emphasize inclusiveness, equality, and respect for others, as well as listening and being open to others' views. The recruitment of participants can vary. For experimental purposes, a convenience sample (of students) can be appropriate, but in order to increase external validity, a representative sample of the population is better. Sometimes open calls through media and advertising campaigns are used, but this, of course, might create a bias amongst participants because of self-selection. Before discussing, participants can be given balanced information about the topic, and they may hear and pose questions to an expert panel (expert panels are always part of the Deliberative Poll). Participants usually fill in a survey before and after taking part in deliberation. These *pre* and *post* surveys are then used to study potential changes in different outcome measurements, such as opinions. Small group discussions are also commonly recorded and transcribed and can thereby also be analysed.

Example: An Experimental Study of Enclave Deliberation and Group Polarization

In this section, we describe what a *controlled experiment in citizen deliberation* might look like. In other words, we look at an experiment that was fully funded and designed by academics and did not have a direct connection to a political process. We organized this experiment in 2012 (Grönlund et al. 2015). We were able to cover the costs through project funding from the Academy of Finland (project no. 251222), and could therefore organize it as a controlled lab-in-the-field experiment. In designing the experiment, we were inspired by the concern, most notably expressed by Cass Sunstein (2002, 2007, 2009), about the increased tendency to only discuss politics in like-minded groups. This 'enclave deliberation' may lead to group thinking with extreme views and an amplification of cognitive errors as a result. When like-minded people discuss among themselves, they reinforce each other's views which prevail in the group at the outset. After discussion, their views move in the direction of the initial bias. This phenomenon is called group polarization. Sunstein (2009, 3) defines group polarization as follows: 'members of a deliberating group usually end up at a more extreme position in the same general direction as their inclinations before deliberation began'. Like-minded discussion may also lead to an amplification of cognitive errors (Sunstein 2007, 80–95, 140–143), meaning that people's false beliefs are strengthened. When it comes to deliberation, the presence of conflicting viewpoints is often regarded as a necessary condition (Thompson 2008, 502). However, the term 'enclave deliberation' has been used to refer to any discussion amongst like-minded people (Sunstein 2002, 2007, 2009). We wanted to test how deliberative norms work

in like-minded groups. Our central research question was: 'Can group polarization be avoided?'

The topic for the deliberative experiment was immigration. The participants' opinions on immigration were measured before and after deliberation with a questionnaire. Based on their baseline views, respondents were divided into two enclaves. Those with negative attitudes to immigration formed a *con* enclave, and those with a positive view on immigration formed a *pro* enclave. Since the main research interest was enclave deliberation, we manipulated the group composition in order to compare deliberation in two types of groups: (1) groups with similar views on immigration; and (2) groups with diverse opinions on immigration. Thus, the participants were randomly assigned to like-minded groups, mixed groups, and a control group. The treatment groups deliberated, whereas the (pseudo) control group only filled in surveys at home.

First, we mailed out a short survey to a simple random sample of 12,000 adults in the region of Turku (Åbo). The response rate to this survey was 25 per cent. It consisted of fourteen items measuring attitudes on immigration. Since the design of the experiment required people with clear views on the immigration issue, we excluded moderates (n = 631), that is, those respondents whose opinions on immigration were close to the median value of the frequency distribution (see Grönlund et al. 2015 for a detailed description). Then, a second survey (T2), which also included an invitation to take part in a deliberation event, was sent out to 2601 people. Half of the invited sample consisted of *pro*- and the other half of anti-immigrants (*con*). We also clarified that only a part of those who volunteered could be included in the deliberation event and that the choice would be made by lot. Each participant in deliberation received a remuneration of 90 euros, whereas the control group received 15 euros.

Ultimately, 805 people volunteered, and we invited 366 people to take part in the deliberative experiment. The target sample was 256 participants, that is, thirty-two small groups of eight participants each. Alas, only 207 people showed up. Especially people in the *con* enclave tended to abstain at this final stage, even though there were no indications of this kind of a bias earlier. For the experiment, we randomly assigned people into treatment conditions and small groups within the *con* and *pro* enclaves. In the end, we were able to form ten *pro* like-minded groups, five *con* like-minded groups, and eleven mixed groups. Because of attrition at the last stage, and the need for balance between the enclaves, each mixed group consisted of exactly eight participants—four from both enclaves—whereas the group size was allowed to vary between seven and nine in the like-minded treatment. The pseudo control group consisted of 369 people. The deliberation event took place during the course of a weekend in the spring of 2012. Each participant took part during one day. The setup was identical. Both days started with a fifteen-item knowledge quiz, after which the participants were briefed in plenum about some basic and balanced facts related to immigration in Finland. The briefing was also given as a handout to the participants.

Deliberation took place in small groups and lasted for four hours, including a forty-five-minute lunch break. Trained moderators facilitated the discussions and made

sure that the discussion rules were followed. First, each participant put forward a theme related to the immigration issue. The discussion started with these issues and was free—the moderators interfered only if any of the group members dominated or completely withdrew from the discussion. After deliberation, each participant filled in a post-test survey.

The main result of the experiment was that all participants in the anti-immigration enclave became more liberal. This was especially true in the mixed treatment, where the groups consisted of four anti- and pro-immigrants each, but also in the like-minded groups consisting of participants with initially negative views on immigration. Within the pro-immigration enclave, participants in the mixed treatment did not change their preferences, whereas a slight polarization into a liberal direction could be traced in the pro like-minded groups. Especially those pro-immigration participants in like-minded groups who did not learn in the course of deliberation became polarized, that is, even more liberal in the post-deliberation measurement (Grönlund et al. 2015). The results support a central theoretical assumption in deliberative democracy, which claims that all opinions should not have an equal weight in the process of public reasoning. Reasonable arguments appealing to generalizable moral principles should be powerful, whereas arguments based on attitudes such as prejudice should be 'laundered' in the course of deliberation (Goodin 1986). Thus, our interpretation of the outcome of the experiment was that deliberation is different from free discussion. The deliberative package with information material and discussion rules emphasizing respect, equality, and reflection can be particularly useful if we want to prevent group polarization in like-minded groups (Grönlund et al. 2015).

The *data* that can be collected through a deliberative mini-public experiment are plentiful. First, survey instruments are good for tracing changes in opinions. Comparisons can be made in two ways: (1) within treatments, using the pre-test, post-test design and measuring the same individuals before and after deliberation; (2) a post-test-only strategy can be used by comparing treatment group(s) with a control group. Both methods have some weaknesses. Using the within treatments method, a phenomenon called 'regression to the mean' might be a problem. Regression to the mean is a phenomenon where measuring the same sample twice—such as in a pre-test, post-test design—tends to lead to regression towards the group's natural mean upon the second measurement, creating an illusion of a treatment effect (Torgerson and Torgerson 2008, 10–15). Using the between treatments method, one must be certain that the randomization has succeeded, and that the treatment and control groups are not different in certain characteristics, such as education, age, or gender. In small samples, this might be an issue. Thus, if possible, a combination of within treatments and between treatments testing can reduce the risk of false findings, especially to reduce type I errors caused by regression to the mean.

Second, data can be generated to analyse the contents and dynamics of deliberation. For this, the discussions need to be at least audio recorded and transcribed. Such endeavours are often time-consuming, at least with rarer languages such as Finnish and Swedish, involving manual transcriptions and coding (Himmelroos 2012;

Himmelroos and Christensen 2014), but they can also be automated, as Kaiping Chen (2019, chapter 2) demonstrates. There is no clear-cut standard for content analysis of deliberation, but the most promising tool for measuring the 'deliberativeness' of discussions is the Discourse Quality Index (DQI), developed by Steenbergen et al. (2003), through an analysis of discussions in the British House of Commons. Another way of analysing deliberation in small groups is to look at the dynamics with the help of network analysis (Gerber 2014): who addresses whom and how? For analysing the dynamics of the behaviour in groups, video recording is another possibility, but the interference of cameras would, of course, present another treatment, which ideally should be controlled for (microphones are less conspicuous than cameras). In addition, the moderators can be surveyed or interviewed afterwards. We have used this method in order to trace social pressures at the small group level, and compared these surveys with similar survey questions to participants (Setälä et al. 2010).

Conclusion

Experiments on deliberative democracy most commonly study *micro-level phenomena*. A typical deliberative experiment asks what kinds of individual-level outcomes certain institutional arrangements have. For example, we can test whether deliberative discussion produces the types of outcomes it is claimed to produce, or study more specifically different ways to organize deliberative discussions and study their consequences. Apart from naturally occurring ones, experiments are not well suited to studying *macro-level phenomena* because control and experimental manipulation are hard to achieve at the macro level. However, naturally occurring experiments allow for a study of macro-level effects. An example is the use of democratic innovations in certain municipalities and not in others (e.g. Touchton and Wampler 2014). Comparing those municipalities, where democratic innovations are used to those where they are not, can also be done with regard to certain macro-level indicators, for example, voter turnout. In principle, it is also possible to study cross-country variation, but differences between democratic systems makes this even harder than comparing variation within a country at the local level.

Controlled experiments are harder to implement at the macro or even meso level. Therefore, experimental research is not likely to give answers to current issues on the systemic nature of deliberative democracy (Parkinson and Mansbridge 2012). We are not able to experimentally manipulate the deliberativeness of a whole democratic system and test the consequences.

Although a large number of experiments with deliberative mini-publics have already been carried out, it is clear that there is room for more experimentation in the field. We separate four areas of interest where we see a need for further experiments, but it is likely that there are other relevant directions for future research. Controlled experiments are typically conducted within a certain context, whereas systematic comparative experimental research, for example, between countries, is lacking. Thus,

it would be important to obtain knowledge of the context dependency of experimental results, and to increase the external validity of the results. Second, there are a number of ways to organize mini-publics, and we lack comparative evidence on the influence of different types of mini-publics on the consequences of deliberation. Third, we also lack knowledge as to how certain individual-level psychological processes interact with participation in deliberative discussion (Mercier and Landemore 2012). We know that people tend to be biased in their information processing, that is, mainly seeking evidence that confirms their existing views (Lodge and Taber 2013), but we do not know whether taking part in deliberative discussion would alleviate these types of biases. And finally, while it seems to create certain problems, it also seems essential to run mini-public experiments that have connections to real-world political decision-making in order to be able to test how being part of a real decision-making process influences the consequences of deliberation, both amongst the participants and in the mini-public as a whole.

Further Reading

Druckman, James, Donald Green, James Kuklinski, and Arthur Lupia. 2011. *Cambridge Handbook of Experimental Political Science*. Cambridge: Cambridge University Press.

Morton, Rebecca, and Kenneth Williams. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab*. Cambridge: Cambridge University Press.

Webster, Murray, and Jane Sell, eds. 2007. *Laboratory Experiments in the Social Sciences*. Burlington: Elsevier.

References

Breckon, Jonathan, Anna Hopkins, and Ben Rickey. 2019. Evidence vs Democracy: How 'Mini-Publics' can Traverse the Gap between Citizens, Experts, and Evidence. Nesta: Alliance for Useful Evidence, available at: https://www.alliance4usefulevidence.org/publication/evidence-vs-democracy/.

Chen, Kaiping. 2019. When People's Voices Matter: Examining Mini-Public Deliberation and Digital Crowdsourcing with Machine Learning Tools. Palo Alto, CA: Stanford University Press, available at: https://purl.stanford.edu/vd984vj3622.

Dryzek, John, and Simon Niemeyer. 2006. 'Reconciling Pluralism and Consensus as Political Ideals'. *American Journal of Political Science 50*: 634–649.

Fishkin, James. 1991. *Democracy and Deliberation: New Directions for Democratic Reform*. New Haven, CT: Yale University Press.

Fishkin, James. 2018. *Democracy when the People Are Thinking: Revitalizing Our Politics Through Public Deliberation*. Oxford: Oxford University Press.

Gerber, Marlène. 2014. 'Equal Partners in Dialogue? Participation Equality in a Transnational Deliberative Poll (Europolis)'. *Political Studies 63* (51): 110–130.

Goodin, Robert E. 1986. 'Laundering Preferences'. In *Foundations of Social Choice Theory*, edited by Jon Elster and Aanund Hylland, pp. 75–101. Cambridge: Cambridge University Press.

Goodin Robert, E., and John S. Dryzek. 2006. 'Deliberative Impacts: The Macro-Political Uptake of Mini-Publics'. *Politics & Society 34* (2): 219–244.

Gosnell, Harold. 1926. 'An Experiment in the Stimulation of Voting'. *American Political Science Review 20* (4): 869–874.

Grönlund, Kimmo. 2016. 'Designing Mini-Publics for Constitutional Deliberative Democracy'. In *Constitutional Deliberative Democracy in Europe*, edited by Min Reuchamps and Jane Suiter, pp. 109–128. Colchester: ECPR Press.

Grönlund, Kimmo, André Bächtiger, and Maija Setälä. 2014. 'Introduction'. In *Deliberative Mini-Publics: Involving Citizens in the Democratic Process*, edited by Kimmo Grönlund, André Bächtiger and Maija Setälä, pp. 1–8. Colchester: ECPR Press.

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

Grönlund, Kimmo, Maija Setälä, and Kaisa Herne. 2010. 'Deliberation and Civic Virtue: Lessons from A Citizen Deliberation Experiment'. *European Political Science Review 2* (1): 95–117.

Himmelroos, Staffan. 2012. *Det Demokratiska Samtalet: En Studie Av Deliberativ Demokrati I Ett Medborgarforum*. Åbo: Åbo Akademis förlag.

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

Lodge, Milton, and Charles Taber. 2013. *The Rationalizing Voter*. New York: Cambridge University Press.

Mansbridge, Jane. 2010. 'Deliberative Polling as the Gold Standard'. *The Good Society 19* (1): 55–62.

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

Morton, Rebecca, and Kenneth Williams. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab*. Cambridge: Cambridge University Press.

Parkinson, John, and Jane Mansbridge. 2012. *Deliberative Systems: Deliberative Democracy at the Large Scale*. Cambridge: Cambridge University Press.

Setälä, Maija, and Kaisa Herne. 2014. 'Normative Theory and Experimental Research in the Study of Deliberative Mini-Publics'. In *Deliberative Mini-Publics: Involving Citizens*

in the Democratic Process, edited by Kimmo Grönlund, André Bächtiger, and Maija Setälä, pp. 59–76. Colchester: ECPR Press.

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.

Steenbergen, Marco, André Bächtiger, Markus Spörndli, and Jürg Steiner. 2003. 'Measuring Political Deliberation: A Discourse Quality Index'. *Comparative European Politics 1* (1): 21–48.

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. 2002. 'The Law of Group Polarization'. *The Journal of Political Philosophy 10* (2): 175–195.

Sunstein, Cass. 2007. *Republic.com 2.0*. Princeton, NJ: Princeton University Press.

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

Thompson, Dennis, F. 2008. 'Deliberative Democratic Theory and Empirical Political Science'. *Annual Review of Political Science 11*: 497–520.

Torgerson, David, and Carole Torgerson. 2008. *Designing Randomised Trials in Health, Education and the Social Sciences: An Introduction*. Basingstoke: Palgrave Macmillan.

Touchton, Michael, and Brian Wampler. 2014. 'Improving Social Well-Being Through New Democratic Institutions'. *Comparative Political Studies 47* (10): 1442–1469.