4-16-2012

Favorable Conditions to Epistemic Validity in Deliberative Experiments: A Methodological Assessment

Didier Caluwaerts
Vrije Universiteit Brussel, dcaluwae@vub.ac.be

Juan E. Ugarriza
El Rosario University - Colombia, juan.ugarriza@urosario.edu.co

Follow this and additional works at: https://www.publicdeliberation.net/jpd

Favorable Conditions to Epistemic Validity in Deliberative Experiments: A Methodological Assessment

Abstract

Methodological evaluations of deliberative mini-publics usually focus on the internal and external validity of experimental designs. Even though such a focus on causal inference and generalization is important, it is incomplete. We argue that the epistemic validity of experimental designs should also be taken into account in order to ensure measuring truly a deliberative exercise rather than just a regular discussion. By ensuring the inclusion and publicity of all arguments, the process of arguing back-and-forth between multiple positions is theoretically claimed to generate better outcomes and should therefore be validated along epistemic lines. Here, we suggest some methodological techniques for enabling the epistemic validity assumption of deliberative experimental designs. These techniques relate to the sampling of the groups and the treatments they receive.

Keywords

deliberative democracy, epistemic proceduralism, experimental methodology, cognitive diversity

Acknowledgements

The steady rise of deliberative democracy on the political firmament initiated numerous attempts to put the deliberative ideal into operational terms. So far, the results are promising and the philosophical assumptions seem to corroborate well. Deliberation is found to instigate more considered judgment (see e.g. Gastil and Dillard 1999), and to reduce preference divergence among deliberants (See e.g. Barabas 2004). However promising these results may be, as deliberative democracy moves from a philosophical ideal to an empirical working theory, scholar's attention should broaden from questions of normativity to issues of validity. Because of the doubts about the viability of genuine deliberation, there is a high demand for methodologically sound demonstration of its existence.

Methodological evaluations of deliberative endeavours mostly boil down to assessments of causality and generalizability. Experiments have to be internally valid, in that they have to be able to demonstrate causal relations through controlling for confounding variables, but at the same time the results have to have some wider application than the sample of citizens that was asked to deliberate. What is often overlooked though is that deliberative experiments have a very specific nature: they are experiments and therefore they have to meet standards of internal and external validity, but they are also experiments with a normative ideal. Every deliberative event, no matter what its empirical or theoretical finality, embodies the idea that decisions made through deliberation generate in some way "better" solutions. According to these normative premises, the representation of a multitude of opinions in a deliberative body ensures that good arguments are identified and bad ones are eliminated (Dryzek 2000). Deliberation thus requires difference in opinion and perspective, and Thompson even argues that "if the participants are mostly like-minded or hold the same views before they enter into the discussion, they are not situated in the circumstances of deliberation" (Thompson 2008, p. 502). As such, deliberative experiments differ from mainstream experiments in that they have an additional requirement: they should be validated along epistemic lines.

This epistemic validity is largely overlooked in the positive theory of deliberation, whereas it features prominently in the philosophical ideal (Estlund 1997; Nino 1996). We therefore ask ourselves how the epistemic benefits that are normatively attributed to deliberation, can be methodologically anchored in empirical research. We argue that epistemic validity can be significantly enhanced by techniques related to the sampling and treatment phases of mini-publics.

We should note in advance that the argument we present here draws heavily on the experimental literature of deliberation, and the approach we take is a scientific one by demanding proof of validity. This might make deliberative practitioners feel uneasy. However, the concern of including participants in

deliberation who are epistemically diverse is a general one for both "action researchers" and scientific experimenters. As such, deliberative practitioners might find the techniques we propose to be of interest for their purposes too.

This paper proceeds as follows. First, we show how epistemic concerns are inherently part of deliberative scholarship. Putting together the principles of publicity and inclusion automatically implies the need for a multitude of contesting opinions. Next, we determine what are the necessary prerequisites of epistemic validity, and what are its consequences. Drawing on the literatures on cognitive diversity and agonistic inquiry, the second paragraph shows the importance of heterogeneity within deliberative groups. Thirdly, we present a list of methodological techniques that can enhance the epistemic attributes of a deliberative experiment, after which we go into detail on how to assess the epistemic qualities of a deliberation. Fifthly, we consider the threats to epistemic validity. And finally, we discuss the relationship between epistemic, and internal and external validity.

Deliberative episteme between publicity and inclusion

The deliberative turn in political philosophy overshadowed the epistemic turn within deliberative theory itself. Many scholars have recently defended the idea that deliberative decision-making procedures are normatively appealing because the outcomes they generate are substantively better than those under different procedures (Lafont 2006). Arguing back-and-forth and weighing reasons pro-and-con allows for perspective taking, and through the cognitive incorporation of other citizens' standpoints, justificatory cross-pressures foster representative thinking (Arendt 2005). The deliberative process is therefore considered to be higher on epistemic validity than other decision-making procedures. By epistemic validity we understand that the outcome of a deliberative decision-making process is considered better because the substance of a decision approximates a reasonable and true solution.

Such epistemically valid deliberation results from a process of arguing that is both inclusive and public (Bohman 1996). Inclusion means that everyone who is subjected to the consequences of a decision should be included in the process leading to the decision. Only when all those concerned have equal opportunities or capacities to participate in a discussion, and only when all actors and their opinions are considered inherently valuable to finding a common solution can a reasonable consensus emerge (Cooke 2000). The reason why inclusion is crucial for the epistemic superiority of deliberation is that all of the problems democracies are faced with are unevenly distributed among the citizenry. Different experiences lead to different perspectives on what constitutes a social or political problem, and this interpretive variation is the key

to solving the problems. After all, if certain problems disproportionately affect certain social groups, these groups should be included in deliberations aimed at overcoming these perverse effects (Bohman 2007; Fearon 1998).

Deliberation also has to be characterized by publicity. Justifications should be formulated so that other participants can understand and can reasonably be expected to accept them. This requirement not only forces the participants to think through their own arguments, but also to always take other participant's sensitivities into consideration when developing and elaborating arguments (Benhabib 1996, p. 71). Publicity guarantees that all arguments are subjected to a wide range of alternative arguments in a situation of communicative symmetry. It helps to make sure that all arguments are tested in a non-coercive and rational discussion.

The joint application of inclusion and publicity means that any deliberation should capture the full range of positions and opinions on an issue of public concern. The process of deliberation thus instigates mental reflection and, under the conditions of the ideal speech situation, the conclusion a deliberative group reaches is a better reflection of the will of the group, and a better solution to the problems faced by the group. Hence, the epistemic quality of deliberation lies at the interplay between dialogical interaction and the internalization of conflicting arguments (Goodin & Niemeyer 2003).

Decisions made under such conditions are claimed to be more legitimate, and of superior quality than those made by aggregation procedures (Cohen 2002). Public reason and consensus can only come about under the condition of maximal inclusion and diversity of opinions. Nevertheless, it should be noted that each consensus is merely the fallible and provisional outcome of the exchange of rational arguments. All arguments and positions must be considered hypothetical. The epistemic qualities of a decision are thus open to revision, and the appearance of new evidence or perspectives can open up the consensus that was reached.

Under what conditions is deliberation epistemically valid?

In order to meet its promises of epistemic superiority, deliberation should take place under conditions that allow for the discursive representation of a broad scope of interpretations of, and perspectives on an issue of public concern. The issue of epistemic validity is therefore highly dependent on the cognitive diversity of the group in which deliberation takes place (Anderson 2006; Landemore 2008). Cognitive diversity generates such a dynamic through two processes. First and foremost, deliberation has epistemic qualities because it involves a process of information pooling. In order to solve problems of a collective nature, participants are required to contribute their pieces to the

puzzle, by giving their own knowledge and experience a public character. As such it is not a mere exchange of facts, but a pooling of perspectives, frames and interpretations (Martì 2006).

Besides information pooling, the process of argumentation also has epistemic merits. Once the veracity of facts and rightness of interpretations is sorted out, arguments in favor of or against certain opinions have to be formulated based on the common pool of information within the group. This process of arguing back and forth has to happen under the conditions of the *ideal speech situation*, i.e. each argument and opinion has to have equal status, and participants have to yield to the "forceless force" of the better argument (Habermas 1981).

Since opinions and perspectives on an issue of public interest are socially stratified, Landemore claims that social groups should be descriptively represented in deliberative mini-publics (Landemore 2010). The deliberating group would thus reflect on a small-scale the interpretive diversity of the larger democracy, and its epistemic potential could be fully explored.

Landemore's argument resonates the "diversity trumps ability"-theorem (Hong & Page 2001; Page 2007). This theorem holds that a diverse group of problem solvers, who are not necessarily the most able, will outperform a homogeneous group of the best problem solvers. It therefore sharply contrasts with the idea that the epistemic validity of deliberation "increases dramatically to the extent that the access to the decision-process is restricted to the wiser" (Martì 2006, p. 48).

Even though he reckons it to be important for epistemic validity, Bächtiger critiques this idea of cognitive diversity (Bächtiger 2010, p. 13). According to him, Landemore's theory sticks too much to the Habermasian ideal type of rational and impassionate discussion, and lacks specific indications of how cognitive diversity translates into the potential for epistemically superior deliberations. He therefore stresses the idea of agonistic inquiry, a process of interaction that is geared to more confrontational modes of interaction.

Two psychological processes figure centrally in bringing about such a form of interaction. The first one is questioning. Rather than merely pooling information and perspectives, agonistic inquiry delivers epistemically superior decisions by critically questioning other deliberants' frames. The aim is not to attack other members personally, but merely to scrutinize the value of others' frames and the importance he or she confers to it. Besides questioning, agonistic inquiry requires disputing, i.e. the process of critically arguing back-and-forth, of weighing arguments and of exacerbating differences in position. This process of questioning and disputing has to be continued until the better arguments claim victory.

Agonistic inquiry is thus more active and adversarial than Habermasian argumentation, but it is just another way of capturing the full ramifications of the arguments and revealing inconsistencies. In this sense, agonistic inquiry mainly attempts to overcome the flaw it sees in ideal-type deliberation. It pushes deliberation to its boundaries by not giving in to an easy fight.

Despite his critique on the lack of process specification, Bächtiger still agrees with the substance of Landemore's claims. Both reckon that cognitive diversity matters to the epistemic quality of deliberative interactions. The question is therefore how cognitive diversity can be methodologically anchored in deliberative experiments in order to promote high epistemic validity.

Methodologically anchoring epistemic validity in deliberative mini-publics

Deliberative democracy is often criticized for being overly optimistic because it is hardly ever possible to guarantee full inclusion and publicity. The epistemic potential of discursive interactions between citizens is thus under threat. This does not have to imply, however, that the design of mini-publics, so-far the main vehicle for advancing deliberative research, cannot be shaped to guarantee that at least the potential for epistemic validity is present. After all, methodological designs can implement some of the favorable preconditions for epistemic validity in the group, all the while acknowledging that "even under good conditions many [decisions] are bound to be incorrect, inferior, or unjust" (Estlund 1997, p. 174).

By proposing procedures that might favor the emergence of epistemic value, we take a procedural perspective to epistemic validity. Some would say that such a procedural approach offers no assurance that the substantive outcomes are in effect better from an instrumental or intrinsic point of view, but a similar critique can be directed to standard procedures enforced to favor internal and external validity. Experimenters often claim for instance that the procedure of randomization leads to high external validity, but they will never be able to show that an entire population put to the same treatment will have the same outcomes. They rely on the procedural characteristics of their designs to make statements about the likelihood of valid causal inference and generalization.

Similarly, in this paper we don't look if the outcome is in effect epistemically superior, because that would involve the use of ontological standards on fairness or rightness, on which even heated philosophical debates proved endless (Martì 2006). Rather, we evaluate how methodological choices shape the conditions favorable to the emergence of epistemic validity. A number of techniques could qualify for that and none of them is new to experimental research. Their novelty lies, however in the fact that they have

never been seen through the lens of epistemic validity. That is, it has never been demonstrated how the methodological choices experimenters make affect epistemic validity.

Two types of techniques can be distinguished relating to the sampling and treatment of experimental groups. Sampling techniques refer to the composition of the mini-publics and guarantee the reflection of the cognitive diversity of the larger population within the setting of the mini-public. We distinguish between randomization, precision matching, and heterogeneity sampling. The treatment techniques, on the other hand, embody the idea that given a certain group composition, specific interventions can be made to ensure that the full spectrum of public positions is captured. We discuss the use of information booklets, experts, and devil's advocacy.

These methodological handles are neither exhaustive nor mutually exclusive, but what both the sampling and treatment techniques have in common, is the idea that all citizens, as participants to public deliberation, should be regarded as offering equally important inputs, a necessary condition for epistemic validity. These procedures thus ensure that every perspective has a good chance of being integrated in the mini-public (sampling techniques), or that in default of such sampled diversity, treatments ensure the presence of multiple perspectives.

1   Sampling

The techniques that fall under the category of sampling methods to ensure epistemic validity, all relate to some element of the process of composing groups. Different techniques are available to provide groups with the cognitive diversity that is needed to generate epistemically superior outcomes. Random sampling is one of them, and received most attention in the literature because of its claims to representativeness. The use of randomization is however highly dependent on the other methodological choices made by the experimenter, such as group size. We therefore also discuss two other techniques that can raise epistemic validity under less than perfect circumstances, namely precision-matching and heterogeneity sampling.

a. Randomization

The first sampling method to ensure that the multitude of public opinions is present in a group is randomization. The idea of selecting a random sample of citizens to join a deliberative mini-public and discuss matters of political interest, finds support in the "diversity trumps ability"-theorem. As Page argues, random selection and assignment of moderately able participants to groups

generates higher epistemic validity than a small group of highly qualified problem solvers, because of their experiential diversity (Page 2007). It should thus come as little surprise that deliberative philosophers hold the idea that random selection is normatively desirable (Bohman 2007, p. 351-352).

Besides guaranteeing that experimental groups are identical in terms of confounders, randomization also ensures sufficient intragroup variation. The best way of reaching this goal is to compose the experimental groups as a mirror image of the population at large (Landemore 2010). The random selection and assignment of participants to groups is considered epistemically superior because randomization provides a representative cross-section of the perspectives and interpretations circulating in society (Ryfe 2005, p. 52).

Randomization can therefore be considered to be functionally up to the task of ensuring cognitive diversity, and the technique is used extensively in Deliberative Polling. Drawing a simple random sample from the population is considered to give each individual in the population an equal chance of being selected for participation. Randomization, Fishkin and his colleagues contend, "produces discussion among people who think and vote differently and would not normally be exposed to one another" (Fishkin et al. 2000, p. 660). Randomization thus fosters cognitive diversity by stimulating completeness and diversity (Fishkin & Farrar 2005); it gives each opinion an equal chance of being included, and the final sample reflects the diversity in opinions that exists in the minds of the public. As such, it avoids the kind of informational inbreeding among participants with similar backgrounds, which undermines epistemic validity.

b. Precision-matching

Even though randomization is normatively considered to be most desirable, its application is limited to large groups. However, many deliberative experiments gather somewhere between five and ten participants at a time. Under such circumstances, it can be difficult to descriptively represent the wide spectrum of opinions and perspectives by using probability samples. This is especially true when the population consists of small minorities, the inclusion of which is highly desirable.

A sampling technique that holds high expectations for improving epistemic validity in smaller groups is precision matching. Rather than assigning the participants to one group or another on a random basis, this technique identifies pairs of similar cases that are either very similar or very diverse and assigns them to the treatment and control conditions (Johnson et al. 2008). This allows the experimenter to control intragroup heterogeneity as well as intergroup homogeneity. In practice, precision-matching requires a researcher

to identify a number of characteristics that (s)he wants to see represented in each group, and then to look for participants who share the same configurations of characteristics. These characteristics can range from socio-demographic variables to political and issue-specific preferences. Each of these participants is then assigned to a group and receives an experimental treatment or not.

Despite its positive results for both internal and epistemic validity, the main disadvantage of precision matching is that the characteristics on which the selection of participants is made are highly subjective. Why do we expect for instance that it is important for the cognitive diversity that women are descriptively represented in the group? The choice of characteristics could seriously harm the idea of procedural impartiality among the multitude of public perspectives (Estlund 1997, p. 195). The characteristics chosen thus have to be extensively justified. This could of course be inspired by theoretical arguments, but caution is necessary when using matching techniques in order not to create an artificial bias.

c. *Heterogeneity sampling*

The last sampling method that could positively influence the epistemic qualities of mini-publics, is heterogeneity sampling. This technique is somewhat less demanding than the two preceding ones, but is more feasible in practice. Heterogeneity sampling does not ensure that pairs of similar cases are present in each group, but merely ensures that a diversity of perspectives is included in each group as a whole. Even though it allows for less experimental control over confounding variables, experimenters meet the essential demand for epistemic validity, namely to have a multitude of political perspectives represented.

One example of this procedure is used in Citizens' Juries, which involve drawing a so-called stratified random sample from the population (Coote & Lenaghan 1997, p. 9). Even though, it is called random, the sampling ensures that some predetermined characteristics of the population at large are represented in the experimental groups. What is interesting, is that the organizers of Citizens' juries claim that the inclusion of participants should not only be based on socio-demographic characteristics, but that also the attitude and opinion distribution within the sample should be proportionally representative of the larger population (Blamey et al. 2000). Valuing cognitive diversity in this way is somewhat artificial but inherent to the small group design, and meets the requirements for epistemic validity.

2 Treatment

Besides sampling methods, there are also techniques that aim to make cognitive diversity the main aim of the experimental treatment. The epistemic validity of the experimental setting is thus guaranteed by deliberate interventions on behalf of the experimentalist. As such, sampling problems leading to too much intragroup homogeneity can be overcome by promoting cognitive diversity through experimental treatments. Suppose e.g. that only a group of highly educated participants shows willingness to partake in a deliberative experiment, the cognitive diversity – and thus the epistemic validity – of the experiment will be rather low, but these sampling biases can be mitigated by a number of techniques. These techniques are (1) the provision of information booklets with briefing materials that captures a wide variety of public positions towards an issue; (2) the availability of experts, which answer questions and provide different perspectives to an issue; and (3) the use of a devil's advocate to draw attention to the voices that are not being heard. It should be noted beforehand, though, that treatment techniques are not full substitutes for sampling techniques; they are merely patches for sampling biases.

a. Information booklets

Providing information booklets is a first way of exposing the participants in a deliberative mini-public to the diversity in frames, perspectives and interpretations present in a given population. These background materials should contain balanced information and arguments pro and con an issue of public concern. The Deliberative Polling initiative considers these briefing materials to be one of its basic ingredients, and usually, a committee representing all parties to an issue carefully screens these booklets, after which they are sent to the participants (Fishkin & Luskin 2005).

The advantage of these briefing materials is that there is virtually no limit to the amount of information that can be reported in them. There is thus enough space to meticulously develop arguments why something should or shouldn't be done. Moreover, information booklets have a very low threshold: they are very accessible for those who want to inform themselves on an issue of public concern but are fearful of being considered ignorant. As a first confrontation with competing ideas, briefing materials take away the strong antagonism that can characterize immediate face-to-face discussion, and that is shown to affect some social groups more than others (Caluwaerts 2012). At the same time, offering background materials stimulates *deliberation within*, i.e. the internalization of conflicting arguments, and the reconsideration of previous

opinions in light of better arguments, which also contributes to epistemic validity (Goodin & Niemeyer 2003).

There is however one important consideration to be taken into account. These materials have to be absolutely neutral, not only substantively, but also in form. On the one hand, they have to be balanced in content, clearly delineating arguments pro and con a certain position, and ensuring that all public positions toward an issue are included. On the other hand, the form of the materials matters too. They have to be written in a language that every social group can understand. If not, exclusionary tendencies are reinforced even before deliberation actually begins.

b. Experts

Another treatment technique enhancing epistemic validity is the possibility to ask questions to experts on the matter under discussion in the experiment. The expert gives a sense of impartiality and knowledgeability to the arguments presented in the deliberation, because they are often called upon to validate the veracity of facts invoked to support arguments during the process of argumentation.

These experts are, however, mostly limited to academics or policy makers (see e.g. Luskin et al. 2002), and therefore rarely capture the full range of positions or sentiments on an issue. What are missing are so-called experience experts, i.e. people who have been affected by certain public choices, who know what it is to have to deal with the consequences of certain decisions. Even though, the perspectives these experience experts bring to the discussion will be considered less "rational", they are absolutely necessary as a complement to the academic and political experts and thus improve epistemic validity. Moreover, the testimonies of those who experienced the impact of certain social or political problems sheds a different light on the stories behind the objective data, and might lower the threshold for low status participants to engage in deliberation.

c. Devil's advocacy

The idea of devil's advocacy has a long history in psychological and management experiments, but despite recent mentioning (Bächtiger 2010; Mercier and Landemore forthcoming) no publications have been found that applied the technique in deliberative mini-publics. If the deliberating group exhibits little heterogeneity, devil's advocacy can prove to be a productive technique exposing the wider diversity in viewpoints (Schweiger et al. 1989). It does so by building decisional conflict into group deliberation, in order to make the process more effective. By formally ensuring that the group has a devil's

advocate, all the assumptions formulated during the discussion will be severely scrutinized. Thinking through opinions this way ensures that invalid arguments are identified and countered, and therefore yields better decisions (Schulz-Hardt et al. 2000).

Devil's advocacy can be implemented in numerous ways, and experimenters should always report which procedure was used. One way of ensuring devil's advocacy is by integrating an accomplice of the moderator in the group who studies all of the perspectives there are to the theme under discussion. This ensures that arguments will be thoroughly questioned and reflected upon. Another advantage is that the experimenter has good control over the treatment as (s)he sets the parameters within which the devil's advocate can manoeuvre. There is, however, more deception at play than in other deliberative experiments, which makes the debriefing in the end all the more important.

The same result can be attained by instructing a subgroup of those partaking in the experiment to take upon themselves the role of devil's advocates. Despite the fact that there are fewer concerns in terms of ethics and artificiality, it could be somewhat harder to bring out the cognitive diversity in the group under such a treatment. Unexpectedly assigning the role to certain participants might take them by surprise, and the lack of preparation might lead to personal attacks rather than the process of taking counter positions (Murrell et al. 1993). If this strategy is pursued, experimenters must at least ensure that the participants get sufficient opportunities to incorporate the different arguments beforehand. Otherwise, there is no added value in terms of epistemic validity.

Assessing the epistemic validity of experimental procedures

After the deliberative mini-publics have taken place, researchers might be interested in assessing whether their experiments were epistemically valid, i.e. whether the cognitive diversity in the group in effect led to the integration of multiple viewpoints. After all, just like experimenters can verify the internal, causal validity of their endeavors with statistical tools, epistemic validity should be assessed ex post. Two approaches stand out as high potential candidates for assessing epistemic validity ex post. The first one aims to adapt the existing measurement instruments to capture the diversity in cognitive perspectives of an issue. During the coding of the deliberative quality of the experiments, the multitude of perspectives can be captured in a number of ways. Mindful of Bächtiger's opinion on cognitive diversity (Bächtiger 2010), however, we take up his suggestion to include some measure of questioning and disputing in the measurements of deliberation. He rightly points to Katharina Holzinger's index,

which contains items pertaining to the epistemic qualities of deliberation, such as the rejection and acceptance of points of view by others (Holzinger 2005). Other categories that might be of use are the formulation of demands, suggestions and questions for clarification.

These adaptations include important coding categories related to cognitive diversity, but they do not determine to what extent there is a genuine exchange of perspectives within the group. The second technique that offers good perspectives for the analysis of the potential for epistemic validity is a frame analysis. A frame analysis has the advantage that it captures the full range of ideas, perspectives and interpretations – i.e. the mental frames people use to structure the world based on proper experiences – in the discussion. Since epistemic validity depends in essence on the discursive representation of different perspectives during deliberation, analyzing the way in which participants frame and reframe their positions is intrinsically valuable (Druckman 2004, p. 674). Such a frame analysis should at least measure how different participants interpret problems and their solutions, what they see as causal mechanisms, and why problems are considered a problem. Capturing these dimensions should yield a comprehensive view of the different perspectives present in the group.

Threats to epistemic validity

The main factors undermining cognitive diversity are processes of self-selection during the recruitment phase and dropout during the organization phase. These should not be problematic as long as they happen randomly within the population or the sample, but wide academic agreement on the stratification of these biases along the lines of gender, class and education raises questions about deliberative scholars' claims to democratic advancement. After all, the socially differentiated engagement in deliberative mini-publics is diametrically opposed to deliberative democrats' aspirations to epistemic validity. If cognitive diversity is to rule deliberative mini-publics, solutions need to be found to ensure the participation of those most in need of inclusive and public decision-making (Bohman 1996).

Even though participants in deliberative mini-publics were found to show their intrinsic motivation and interest after the experiments, ex-post motivation won't get them to attend. Some techniques are available to minimize this risk. Research on response rates in surveys advises to stimulate and motivate the participants through financial stimuli. In any event, the participants will have to be financially compensated for their travel costs and the time they are willing to devote. Moreover, a systematic method of sending reminders was found to have

a positive impact on the attendance rates in focus group research (Krueger 1998).

Despite all efforts, some kind of selection bias is inevitable, unfortunately. Hansen for instance reports on the self-selection and dropout rates along the way of recruiting the participants for the Danish Deliberative Poll (Hansen 2004). More than 40% of those contacted refused to take the questionnaire, and of those remaining, less than half was willing to participate in the poll. Afterwards, there was some additional dropout in the days before the experiment. What is worrying about these figures is that they were strongly stratified along the lines of gender and education. Extra efforts should be put into over recruiting women and people with lower educational attainment.

How does epistemic validity relate to external and internal validity?

Even though epistemic validity should be taken into account as a separate dimension, it is intrinsically related to internal and external validity. However, the relationship between dimensions of validity is not always straightforward, nor linearly positive. It is, after all, often argued that there is actually a trade-off between internal and external validity (Aronson et al. 1995). High levels of external validity ensure generalizability, but they may limit internal validity because randomized selection and assignment of participants to groups, limits the control the experimenter has over confounding variables; it limits the certainty of the causal relation. Conversely, to meet the high demands of internal validity, the experimenter may wish to use participants that are very similar in some way, that are unlikely to drop-out, and conduct the experiment in a perfectly controlled laboratory environment (McDermott 2002).

Now that epistemic validity enters the equation, things become more complex, yet more interesting. The relation between epistemic and internal validity can be considered positive. By systematically ensuring that cognitive diversity is an attribute in all experimental conditions, techniques enhancing epistemic validity will also raise internal validity. By keeping constant the internal heterogeneity of each group partaking in an experiment, confounding variables can be controlled, which in turn makes for strong causal inferences.

The relationship between epistemic and external validity is mixed. In large groups, randomization ensures that results are generalizable, while at the same time ensuring that the mini-public is a mirror of the larger diversity in society. In small groups, randomization is not an option, so that other sampling techniques are used. This could be potentially beneficial for epistemic validity since experimenters have the option of actively assuring intragroup heterogeneity. However, the artificiality of oversampling minority voices for

instance limits the external validity. Hence, in small group experiments, epistemic and external validity are in a trade-off.

Conceptualized as such, the distinction between the three types of validity is clear. External validity refers to the overlap or congruence between sample and population. Internal validity is mainly a function of intergroup homogeneity. That is to say that control for confounders is higher when the experimental groups are more comparable. Epistemic validity, finally, deals with intragroup heterogeneity, with the fact that each group represents the multitude of political perspectives and public opinions that characterizes the citizenry in diverse polities.

Moreover, threats to epistemic validity also plague internal and external validity. The two main problems we identified, self-selection and dropout, potentially limit cognitive diversity, but they also affect internal and external validity. The systematic dropout of the least educated, for instance, limits the generalizability of the experimental findings to the entire population, and when participants drop out in the treatment group, but not in the control group, it is difficult to make causal inferences.

Conclusion

Deliberative theory has taken an epistemic turn in the last few years, but the philosophical argument has only scarcely found its way to the real world of deliberative experimental methodology. Nevertheless, epistemic validity should be an essential concern to anyone interested in experiments with citizens in deliberative mini-publics. After all, the question whether deliberative decision making yields epistemically superior outcomes is founded in the deliberative principles of inclusion of all perspectives on an issue of concern, and publicity in justification.

Moreover, the epistemic validity of deliberative experiments can only be guaranteed when the full spectrum of opinions is reflected in the deliberative group. This cognitive diversity ensures that all those involved have a say and can add perspectives and interpretations to the discussion. What is important for the organization of deliberative experiments, is the discursive representation of the perspectives of all those affected by an issue of public concern. To ensure this cognitive diversity, we've taken a procedural approach to epistemic validity by offering some sampling and treatment techniques that create favorable conditions for its emergence.

The sampling techniques involved randomization, precision matching and heterogeneity sampling. All three attempt to make the multitude of perspectives an inherent part of the group composition, but in different ways. Randomization is based on probability samples of the population and gives everyone an equal

chance of being selected and representing his or her perspective. Randomization is however difficult in small group settings, which induces the need for alternatives. Those alternatives are precision matching and heterogeneity sampling, which require the experimenter to choose a number of characteristics (s)he wants represented in the mini-public, and select participants on that basis.

Whenever sampling techniques don't succeed, the experimenter can bring interpretive diversity into the equation by submitting the participants to certain treatments. These treatments share the idea that active interventions on behalf of the experimenter will lead to the discursive representation of the multitude of perspectives in the mini-public. This can be done through providing briefing materials, questioning experts and incorporating a devil's advocate in each group.

As a methodological ideal for deliberative mini-publics, experimenters should not only ensure epistemic validity ex ante. They should also evaluate it ex post. We argued that two analytical techniques are available to this end. On the one hand, (s)he could adapt his or her measurement instrument by including items that capture processes of questioning and scrutinizing. These processes aim at discovering the diversity within the group. On the other hand, the experimenter could perform a frame analysis, deconstructing the mental frames and interpretations the participants have of the issue under discussion. As such, the frame analysis could reveal the discursive representation of the diverse opinions within the group, as well as processes of frame congruence.

Epistemic validity is an inherent part of experimentation in deliberative scholarship, just like its internal and external counterparts. After all, the potential for generating epistemically superior outcomes is one of the most widely acclaimed benefits of democratic deliberation. Therefore experimenters should put considerable effort into guaranteeing the epistemic validity of their research designs all the while balancing it against the need for internal and external validity. Only deliberative experiments that acknowledge the epistemic powers of citizen discourse can generate high quality results verifying or falsifying the theoretical assumptions underlying deliberative theory.

References

Anderson, Elizabeth. 2006. The Epistemology of Democracy. *Episteme: A Journal of Social Epistemology*, 3, 8-22.

Arendt, Hannah. 2005. Truth and Politics. Pp. 295-314 in José Medina and David Wood (eds), *Truth. Engagements across Philosophical Traditions*. Oxford: Blackwell.

Aronson, Elliot, Phoebe C. Ellsworth, J. Merrill Carlsmith, and Marti Hope Gonzales. 1995. *Methods of Research in Social Psychology*. New York: McGraw-Hill.

Bächtiger, André. 2010. On Perfecting the Deliberative Process: Agonistic Inquiry as a Key Deliberative Technique. Paper prepared for delivery at the 2010 Annual Meeting of the American Political Science Association, Washington (DC), September 2010.

Barabas, Jason. 2004. How Deliberation Affects Policy Opinions. *American Political Science Review*, 98, 687-700.

Benhabib, Seyla. 1996. Toward a Deliberative Model of Democratic Legitimacy. Pp. 67-94 in Seyla Benhabib (ed), *Democracy and Difference*. Princeton: Princeton University Press.

Blamey, Russell K., Rosemary F. James, R. Smith, and Simon J. Niemeyer. 2000. *Citizens' Juries and Environmental Value Assessment*. Canberra: Australian National University.

Bohman, James. 1996. *Public Deliberation. Pluralism, Complexity and Democracy*. Cambridge: The MIT Press.

Bohman, James. 2007. Political Communication and the Epistemic Value of Diversity: Deliberation and Legitimation in Media Societies. *Communication Theory*, 17, 348-55.

Caluwaerts, Didier. 2012. Waar en wanneer spreken mannen en vrouwen over politiek? De sekseverschillen in politieke discussie in hun sociale en politieke context. *Res Publica* 53, 3-23.

Cohen, Joshua. 2002. Deliberation and Democratic Legitimacy. Pp. 87-106 in David Estlund (ed), *Democracy*. Oxford: Blackwell.

Cooke, Maeve. 2000. Five Arguments for Deliberative Democracy. *Political Studies*, 48, 947-69.

Coote, Anna and Jo Lenaghan. 1997. *Citizens' Juries: Theory into Practice*. London: Institute for Public Policy Research.

Druckman, James N. 2004. Political Preference Formation: Competition, Deliberation, and the (Ir)relevance of Framing Effects. *American Political Science Review*, 98, 674-86.

Dryzek, John S. 2000. *Deliberative Democracy and Beyond. Liberals, Critics, Contestations*. Oxford: Oxford University Press.

Estlund, David. 1997. Beyond Fairness and Deliberation: The Epistemic Dimension of Democratic Authority. Pp. 173-204 in James Bohman and William Rehg (eds), *Deliberative Democracy: Essays on Reason and Politics*. Cambridge: MIT Press.

Fearon, James D. 1998. Deliberation as Discussion. Pp. 44-68 in Jon Elster (ed) *Deliberative Democracy*. Cambridge: Cambridge University Press.

Fishkin, James S. and Cynthia Farrar. 2005. Deliberative Polling: From Experiment to Community Resource. Pp. 68-79 in John Gastil and Peter Levine (eds), *The Deliberative Democracy Handbook: Strategies for Effective Civic Engagement in the 21st Century*. San Francisco: Jossey-Bass.

Fishkin, James S. and Robert C. Luskin. 2005. Experimenting with a Democratic Ideal: Deliberative Polling and Public Opinion. *Acta Politica*, 40, 284-98.

Fishkin, James S., Robert C. Luskin, and Roger Jowell. (2000). Deliberative Polling and Public Consultation. *Parliamentary Affairs*, 53, 657-66.

Gastil, John and James P. Dillard. 1999. Increasing Political Sophistication Through Public Deliberation. *Political Communication*, 16, 3-23.

Goodin, Robert E. and Simon J. Niemeyer. 2003. When Does Deliberation Begin? Internal Reflection versus Public Discussion in Deliberative Democracy. *Political Studies*, 51, 627-49.

Habermas, Jürgen. 1981. *Theorie des kommunikativen Handelns*. Frankfurt am Main: Suhrkamp.

Hansen, Kasper M. 2004. *Deliberative Democracy and Opinion Formation*. Odense: University of Southern Denmark.

Holzinger, Katharina. 2005. Context or Conflict Types: Which Determines the Selection of Communication Mode. *Acta Politica*, 40, 239-254.

Hong, Lu and Scott E. Page. 2001. Problem Solving by Heterogeneous Agents. *Journal of Economic Theory*, 97, 123-63.

Johnson, Janet B., Henry T. Reynolds, and Jason D. Mycoff. 2008. *Political Science Research Methods*. Washington: Congressional Quarterly Press.

Krueger, Richard A. 1998. *Moderating Focus Groups*. Thousand Oaks: Sage.

Lafont, Cristina. 2006. Is the Ideal of a Deliberative Democracy Coherent? Pp. 3-26 in Samantha Besson and Jose L. Marti (eds), *Deliberative Democracy and its Discontents*. Aldershot: Ashgate.

Landemore, Hélène. 2008. How To Make Democracy Smarter: on the Proper Use of Deliberation, Majority Rule, and Information Markets for Intelligent Democratic Decision-Making. Paper prepared for presentation at the symposium "Making Democracy Smarter: Markets or Deliberation?", September 2008.

Landemore, Hélène. 2010. Deliberation and the Epistemic Function of Parliamentary Assemblies: An Argument for Descriptive Representation. Paper presented at APSA Annual Convention, Washington, September 2010.

Luskin, Robert C., James S. Fishkin and Roger Jowell. 2002. Considered Opinions: Deliberative Polling in Britain. *British Journal of Political Science*, 32, 455-78.

Martì, Jose L. 2006. The Epistemic Conception of Deliberative Democracy Defended. Reasons, Rightness and Equal Political Autonomy. Pp. 27-56 in Samantha Besson and Jose L. Marti (eds), *Deliberative Democracy and its Discontents*. Aldershot: Ashgate.

McDermott, Rose. 2002. Experimental Methodology in Political Science. *Political Analysis*, 10, 325-42.

Mercier, Hugo and Hélène Landemore. Forthcoming. Reasoning is for Arguing: Understanding the successes and failures of deliberation. *Political Psychology*.

Murrell, Audrey J., Alice C. Stewart, and Brent T. Engel. 1993. Consensus Versus Devil's Advocacy: The Influence of Decision Process and Task Structure on Strategic Decision Making. *Journal of Business Communication*, 30, 399-414.

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

Page, Scott E. 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools and Societies*. Princeton: Princeton University Press.

Ryfe, David M. 2005. Does Deliberative Democracy Work? *Annual Review of Political Science*, 8, 49-71.

Schulz-Hardt, Stefan, Dieter Frey, Carsten Lüthgens, and Serge Moscovici. 2000. Biased Information Search in Group Decision Making. *Journal of Personality and Social Psychology*, 78, 655-69.

Schweiger, David M., William R. Sandberg, and Paula L. Rechner. 1989. Experiential Effects of Dialectical Inquiry, Devil's Advocacy, and Consensus Approaches to Strategic Decision Making. *Academy of Management Journal*, 32, 745-72.

Thompson, Dennis F. 2008. Deliberative democratic theory and empirical political science. *Annual Review of Political Science*, 11, 497–520.