

4-20-2017

Discourse Quality in Deliberative Citizen Forums – A Comparison of Four Deliberative Mini-publics

Staffan Himmelroos

Åbo Akademi University, shimmelr@abo.fi

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

 Part of the [Comparative Politics Commons](#), [Models and Methods Commons](#), and the [Political Theory Commons](#)

Recommended Citation

Himmelroos, Staffan (2017) "Discourse Quality in Deliberative Citizen Forums – A Comparison of Four Deliberative Mini-publics," *Journal of Public Deliberation*: Vol. 13 : Iss. 1 , Article 3.

Available at: <https://www.publicdeliberation.net/jpd/vol13/iss1/art3>

This Article is brought to you for free and open access by Public Deliberation. It has been accepted for inclusion in Journal of Public Deliberation by an authorized editor of Public Deliberation.

Discourse Quality in Deliberative Citizen Forums – A Comparison of Four Deliberative Mini-publics

Abstract

In recent years, there has been a rapid growth in studies reporting findings from a variety of deliberative citizen forums. Such studies help to develop our understanding of deliberative democracy by exploring changes in opinion and knowledge as well as - more recently - the quality of the deliberative process itself. However, most deliberative forums are organized on an *ad hoc* basis, making it hard to judge how generalizable the findings from such forums actually are. This article attempts to address this problem by comparing the findings on the quality of deliberation from four different citizen forums. Based on the findings citizen deliberation is generally very respectful, while argumentation is less refined than among elected representatives. The cases included in this study also suggest that women and those with lower education have less influence in the deliberative process.

Author Biography

Staffan Himmelroos is a postdoctoral researcher at the Social Science Research Institute at Åbo Akademi University. His research is focused on democratic innovations and political behavior. He has published on these topics in *Political Studies*, *International Political Science Review*, *Scandinavian Political Studies* and *Information Polity* among others.

Keywords

Citizen deliberation, Discourse quality, Mini-publics, Comparative research

Acknowledgements

This research was funded by the Academy of Finland (28200124K1)

Introduction

Since democratic theory took a deliberative turn in the early 1990s (Dryzek, 2002), there has been a growing interest in deliberative political practices where citizens resolve their differences through talking rather than voting. It has even been suggested that the deliberative ideal of a reasoned, respectful and open-minded argumentation is most likely to be attained in carefully designed citizen forums, like deliberative mini-publics (Fishkin, 1997; Fung, 2007). To gauge the promise of these deliberative mini-publics, researchers have looked into how policy opinions change as a result of taking part in deliberations (Hansen & Andersen, 2004; Himmelroos & Christensen, 2014; Luskin, Fishkin, & Jowell, 2002; Setälä, Grönlund, & Herne, 2010) and how deliberation might empower the participants (Andersen & Hansen, 2007; Gastil & Dillard, 1999; Grönlund, Setälä, & Herne, 2010; Nabatchi, 2010).

There has, nevertheless, been a lack of studies examining the quality of deliberative process (De Vries et al., 2010; Ryfe, 2005). Considering deliberative democracy's emphasis on the quality of the process by which we reach a decision, one would think that standards of rationality, respectfulness and reflectiveness would be at least as important as the outcome. The good news is that we are seeing a growing number of studies looking into the content of discussions at deliberative citizen forums (Caluwaerts, 2012; Dutwin, 2003; Marlène Gerber, 2015; Karpowitz, Mendelberg, & Shaker, 2012). However, like most studies of deliberative forums, they still tend to focus on a single case. Consequently, we do not know the extent to which the findings from these forums are generalizable, and our understanding of the potential for citizen deliberation will remain limited unless we engage in more comparative work.

The aim of this study is to examine deliberative quality in a more systematic manner by comparing findings from four deliberative citizen forums, which have been analyzed with the help of the same measure, the Discourse Quality Index (DQI). The DQI is an established content analytical measure designed by Steiner et al. (2004) for capturing the quality of deliberative processes. The quality of citizen deliberation is evaluated by (a) *the extent to which citizen deliberation fulfills a number of vital characteristics* ascribed by deliberative theory and (b) *whether deliberative behavior is equally distributed among the participants*. Since there are relatively few studies looking at the content of deliberation, it introduces certain limitations with regard to comparable data. The cases included in this study (a Finnish deliberative experiment; Europolis – a Europe-wide deliberative poll; a deliberative experiment in Belgium; and a deliberative citizen forum in the United

States) all have a similar design but have been arranged in different countries and focus on different topics.

The study is organized in the following manner. First, the deliberative process and its central elements are discussed from a normative standpoint, followed by a critique aimed at the claims forwarded by normative theory. Second, the benefits of more systematic and comparative research in the field of deliberative democracy are discussed. Third, the data from the Finnish deliberative experiment and the three deliberative forums compared to it are described. Thereafter, a description is given of the methods used and the empirical analysis, wherein the different forums are compared to each other. Finally, conclusions based on the analyses are presented. The main finding is that although the deliberative citizen forums show some variation, there are visible similarities. All cases display relatively moderate levels of justification and high levels of respect. With regard to participatory equality, the deliberative forums display concerning signs of inequality as women and those with less education are less active in the discussions. However, the evidence as to whether these groups actually produce lower-quality arguments is inconclusive. Interestingly, younger participants seem to produce higher discourse quality in all of the cases where it was analyzed.

What Defines High-Quality Deliberation?

To assess the quality of deliberation, we need to identify the elements that constitute a democratic deliberative process. According to the theory of deliberative democracy, individuals taking part in deliberation are expected to carefully weigh reasons and exchange morally justifiable arguments in a context of mutual respect (Bohman, 1996; Chambers, 1996; Cohen, 1989; Dryzek, 1990; Gutmann & Thompson, 1996). While political theorists have fairly little to say on what this means in practice, it would appear that a deliberative process should involve at least four basic elements: claims supported by well-defined justifications; concern for the common good; respect for others; and a willingness to consider alternative views. To begin with, if conclusions are to be drawn based on evidence presented in a deliberative process, the claims should be both logical and coherent, i.e. there needs to be an apparent link between presumption and conclusion (Burkhalter, Gastil, & Kelshaw, 2002). In cases where the arguments are longwinded or there is no clear connection between presumption and conclusion, it will be difficult for an audience to evaluate the virtues of the argument (Steenbergen, Bächtiger, Spörndli, & Steiner, 2003).

However, ideal deliberation requires more than the logical justification of opinions and claims; the arguments should also have intrinsic characteristics that make them

compelling to others (Cohen, 1989). Bohman (1996) argues that in order to convince those who disagree with you, you need to understand their reasons and make your reasons understandable to them. Therefore, the arguments presented should consider the well-being of others and of the community at large.

Moreover, a reasonable person should not only have sound reasons to support their opinions; they should also be open to suggestions, open to diverging viewpoints and willing to reevaluate their opinions in the light of new evidence (Chambers, 1996). Everyone involved in the deliberative process should reflect on what is being said and evaluate each argument and the way it relates to their own opinions (Burkhalter et al., 2002; Gastil, 1993). Being open to other opinions and having them influence one's own views requires a fundamental respect for the other participants and their arguments. Young (2002) emphasizes that without respectful consideration for other people's arguments, a deliberative process can never bridge different views or positions that must act alongside each other in every pluralist society. If participants present coherent and logical arguments for their position while simultaneously using derogatory terms or inflammatory rhetoric, the indication is that they are not very interested in persuading their opponents or finding agreement (Chambers, 1996).

For the deliberative process to be considered democratic, it should also guarantee an equal opportunity of access to political influence for all those concerned (Knight & Johnson, 1997). To ensure that an individual's assent to arguments advanced by others is un-coerced, the deliberations should meet the principle of non-tyranny (Fishkin, 1991). Non-tyranny implies that decisions actually reflect the deliberative process, that no group automatically succeeds, and that no group or individual is unfairly disadvantaged in the democratic process by deficiencies due to conditions or circumstances beyond their control (Knight & Johnson, 1997). A fair and inclusive process would subsequently be one where all participants actively take part in the exchange and evaluation of reasoned arguments.

Many scholars have been skeptical as to whether real-world deliberations can be compared to the ideal outlined above (cf. Posner, 2003; Przeworski, 2010). They claim that the bar for a good deliberative process has been set too high and that some individuals are likely to be better endowed than others when engaging in rational argumentation on political issues. The Habermasian ideal speech situation (Habermas, 1984), wherein the normative ideals of deliberation are expected to prevail, has been criticized (see Kohn, 2000) for relying on an assumption that language is fully transparent, i.e. its meaning is accessible to all. As Dryzek & Niemeyer (2008) point out, discourses do not only enable thoughts, speech, and action, they may also constrain them. Every discourse embodies some conception

of common sense and acceptable knowledge, and thus, it may act as an expression of power by recognizing some interests as valid while repressing others. Thus, it has been suggested that the Habermasian take on rational discourse represents a form of communication that is characteristic of some groups while at the same time excluding others (Bohman, 1996; Young, 2002). Sanders (1997) contends that those who are less likely to present their arguments according to the Habermasian ideals are the same people who are already underrepresented and systematically disadvantaged in formal political institutions, namely women, racial minorities and the less educated.

However, it has been suggested that many of the problems associated with citizen deliberation can be mitigated in well-designed deliberative forums (Farrar, Green, Green, Nickerson, & Shewfelt, 2009; Smith, 2009). Organized citizen forums such as deliberative mini-publics — wherein a representative sample of ordinary citizens meet face-to-face in facilitated small groups — are designed to provide favorable conditions for citizen deliberation (Smith, 2009). The participants receive additional information on the issue, and facilitators ensure that the participants can take part in the discussion on equal terms. Hence, deliberative mini-publics are considered to be something of a most likely case, an environment whereby ordinary citizens have the greatest probability of attaining the ideals of deliberative democracy (see Marlène Gerber, 2015). Since mini-publics are created to ensure fair and high-quality deliberation, they also act as a rigorous test of the theoretical assumptions. If citizen deliberation is unable to flourish within such designs, it is unlikely to succeed elsewhere.

Comparing Deliberative Forums

As the evidence base on democratic innovations continues to grow, so does the need for more comparative work (Geissel & Newton, 2012; Ryan, 2014). Comparative studies are imperative if we are to better understand the conditions under which citizen deliberation produces the intended results. Despite some development toward comparative studies of democratic innovations (Carson, 2006; Karlsson, 2010; Ryan & Smith, 2014), little has been done to compare the deliberative exchanges lying at the heart of many of these innovations (see Steiner, 2012, for an exception).

The data used as the baseline for this study is from a Finnish experimental deliberative forum, wherein 135 citizens from a random sample met in small groups to discuss the future of nuclear power in Finland. Since this study focuses on how the findings are representative of such deliberative forums, the Finnish experiment is subsequently compared to other forums that also rely on a diverse, preferably

representative sample, where participants met face-to-face in small groups. The quality of deliberation was measured in the analysis of the Finnish experiment with the help of The Discourse Quality Index (DQI) by Steiner et al. (2004). To be able to make comparisons between different deliberative forums, it was necessary to find comparable data on the quality of the deliberative process from the different deliberative forums. Although there are other ways to measure the quality of deliberative processes (Dutwin, 2003; Holzinger, 2004; Stromer-Galley, 2007), the DQI is probably the measure that has seen the most widespread use. Hence, it should present the best possibility for finding comparable data in a relatively new and little-studied area of research.

Three cases fitting the criteria outlined above were selected. The first case was a pan-European deliberative opinion poll, the second a deliberative experiment arranged in Belgium and the third a citizen forum on broadband organized in the state of Kansas, USA. All three had designs resembling that of the Finnish deliberative forum and had used indicators from the DQI to measure the quality of deliberation. In line with the primary aim of this study, the focus of interest will be on the presence of the different elements of the DQI in the discussions and how the capacity for deliberation is distributed among the participants.

I. Finnish deliberative experiment

The Finnish data originate from an experimental deliberative mini-public arranged in November 2006. The design of the citizen deliberation experiment was similar to a deliberative opinion poll, where a random sample of citizens engages in facilitated small-group discussions (Fishkin, 2009; Luskin et al., 2002). The topic of the discussions was nuclear power, or, more specifically, the participants were asked to make a decision on the question, “Should a sixth nuclear power plant be built in Finland?” Nuclear power was deemed to be a good topic for discussion because it is a salient issue in Finland that many can relate to. The citizen deliberation experiment began by forming a random sample of 2,500 adults from the Turku region in southwest Finland. The final target sample for the experiment was 144 people, i.e. 12 small groups consisting of 12 participants each. Of the invited, 135 participants showed up.

All 12 of the small-group discussions were recorded. However, due to technical challenges (varying audio quality, two recordings failing at different points, etc.) only eight of the small groups could produce transcriptions at the required level of detail, i.e. captured speeches that could be tied to the participant in question by voice recognition. The data thus analyzed comprised eight small groups with 90 participants. The citizen deliberation experiment involved a comparison of two decision-making methods (vote vs. consensus). The data analyzed here represents

the part of the discussion before the decision-making commenced, around three hours of discussion where all groups had a matching treatment (1,189 arguments altogether). Moreover, the participants filled out surveys both before and after the discussions. These surveys measured opinions on energy policy issues and a number of background variables.

Random sampling and group allocation were used to bring about an inclusive process in which all relevant views are represented. Furthermore, the participants received a balanced information package that contributed to their deliberative capacity by equipping them with a basic command of the issue at hand. Before taking part in facilitated group discussions, the participants also met with experts representing different interests to help them in their search for additional information and arguments. In the small-group discussions, each participant was asked to come up with an issue related to the main topic, which would subsequently be used as part of a common agenda upon which the discussion was to be based. They were encouraged to be respectful and attentive toward other participants. Trained facilitators overseeing the process were instructed to intervene only if the discussion halted or to encourage less active participants to speak up. The intention was to generate a free-flowing, constructive, deliberative environment with a fair and balanced discussion (see Setälä et al., 2010, for detailed description of the experiment).

II. Europolis – Deliberative opinion poll

The first citizen forum used to compare the findings from the Finnish deliberative forum was a pan-European deliberative opinion poll known as Europolis, which took place in Brussels in May 2009. It gathered a representative sample of Europeans from all EU member states to deliberate on the issues of migration and climate change (Fishkin, 2009). Following the practice of deliberative opinion polls (Fishkin, 2009; Luskin et al., 2002), opinions were measured before and after the event, balanced reading material was provided to the participants, and they also had the opportunity to discuss the issues with experts and politicians in a plenary session. Altogether, about 350 participants from around Europe took part in the event. The discourse quality has been analyzed in a subsample of 13 small groups from the Europolis deliberative opinion poll (Steiner, 2012; Gerber, 2015).

III. Deliberative Experiment in Belgium

The second citizen forum that was compared to the Finnish experiment was a deliberative experiment arranged in Brussels in 2010. This experiment was run by Didier Caluwaerts and focused on the linguistic cleavage in Belgium, and the primary issue discussed was how the participants see the future of Belgium (Caluwaerts, 2012). Another important element of the experiment was a

comparison of different group compositions and their effects on the deliberations. The selection and assignment of the participants to the experimental groups were based on attitudes measured in a pre-deliberation survey. Altogether, 83 participants attended the deliberative experiment in Brussels. All the discussions were transcribed and analyzed with DQI by Caluwaerts (2012).

IV. Kansas Citizen Deliberation

The third citizen deliberation forum that was compared to the Finnish experiment was organized in different public libraries in the state of Kansas (Han, Schenck-Hamlin, & Schenck-Hamlin, 2015). All in all, 142 individuals took part in small-group discussions on the issue of broadband access. The participants were recruited through an array of different channels. Some were recruited via announcements in newspapers and local radio stations, others via social networking sites, e.g. Facebook. The libraries also contacted their local chambers of commerce and schools to solicit participation. The participants were randomly assigned to 25 small groups, and the discussion lasted for little more than one hour. A total of 23 group discussions were successfully audio-recorded, transcribed and analyzed with the help of DQI (Han et al., 2015).

While, the three deliberative citizen forums I compare to the Finnish experiment have a similar design, it is important to notice that they have been arranged in different countries with different topics. This will naturally introduce some limitations with regard to the comparisons. The research design perhaps best matches what Anckar (2008) defines as a loose most-similar-systems design, where we choose to study cases that appear to be similar in as many background characteristics as possible, but where there is no systematic matching of all relevant control variables. The choice of research design is naturally driven by the quality and quantity of information that is currently available. Nonetheless, the available data represents what we know at a given point in time and can still help us identify mechanisms at play in deliberation and generate hypotheses for further research (Gerring, 2009).

The goal of this study is to identify dominating patterns or mechanisms inherent to citizen deliberation. To help differentiate between variation between the cases and dominating patterns, I contrast the findings of the deliberative citizen forums against parliamentary debates analyzed by Steiner et al. (2004). By using the DQI scores from the Steiner et al. (2004) study on parliamentary debates as a reference point for the deliberative quality, I should be able to identify general patterns of discourse quality relevant to citizen deliberation.

Indicators of Deliberative Quality

Measuring discourse quality

The DQI draws mainly on Habermasian discourse ethics (Steiner et al., 2004). As such, it has strong and apparent connections to the core of deliberative theory. However, it manages to combine the theoretical strengths with a functional outlook. The DQI relies on an idea that deliberative actions can be placed on a continuum from weak deliberation — with insufficient justifications and disrespectful comments — to ideal deliberation — with sophisticated justifications and respectful reciprocal communication. Each speech presented in a deliberative process can be placed anywhere on this scale and thus provide us with an understanding of how close the discussion is to ideal deliberation. The higher a speech act scores on the discourse-quality indicators, the closer it is to the ideal of communicative rationality (Steiner et al., 2004).

Coding for discourse quality begins by making a distinction between speech acts that include a demand and speech acts that lack this feature. A demand is a proposal on what should (or should not) be done. When it has been established that a speech includes a demand, it is evaluated according to a number of predefined categories (see Steiner et al., 2004). Since the different elements of discourse quality do not indicate how active participants are within the small group, the number of speeches involving a demand is used to measure participatory equality and inclusiveness. The *number of arguments* is a purely quantitative measure that denotes the number of arguments each participant produced during the proceedings. This provides a measure of how involved a participant was in the deliberative process.

The original DQI measure included seven indicators (Steiner et al., 2004), but most studies using it have either expanded the number of indicators or decided to concentrate on only a few (Caluwaerts, 2012; Gerber, Bachtiger, Fiket, Steenbergen, & Steiner, 2014). In the theoretical section, it was argued that there are four basic dimensions to the deliberative process: reason-based claims, concern for the common good, respect for others, and a willingness to understand other views. The aim is to include indicators representing each of the four basic dimensions of deliberation. Hence, a slightly modified version of the DQI (with four indicators) was used, which was designed especially for the particular demands of citizen deliberation.

(I) Level of Justification is a measure of how rational the arguments presented by the participants are. The tighter the connection between premises and conclusions, the more rational the justification is and the more useful it will be for deliberation. Qualified and sophisticated arguments include at least one apparent connection

between premise and conclusion, while those with no or inferior justifications completely lack or have inadequate connections. By looking at whether arguments are expressed in the terms of the common good, *(II) Content of Justification* aims to capture the deliberative idea of arguments not only being logical and coherent but also having a general appeal. This indicator aims to capture the motives for an argument. Are the participants looking out for their own interests or are they considering the interests of other people as well? Appeals to the common good can take different forms. The indicator captures both instances when common good is stated in utilitarian terms, i.e. as the best solution for the greatest number of people, and when it is expressed through the difference principle in the sense that the common good is best served if the least advantaged are helped.

The indicator for *(III) Respect* is relatively simple, involving only two categories, explicit disrespect on the one hand and implicit or explicit respect on the other. It measures both how considerate participants are toward other members in the small group and attitudes they reveal toward groups and individuals under discussion. *(IV) Reciprocity* is used to measure how participants react to arguments that contradict their own (Steiner et al., 2004). Are they willing to engage with participants' arguments, and, more importantly, do they respond to counter-arguments? It is also used to identify whether the participants act in a reciprocal manner by weighing or comparing different demands. The coding scheme can be found in Appendix A.

Although these four indicators largely match each of the basic elements presented in the theoretical section, a factor analysis shows that the indicators load on two and not on one dimension (see Appendix C). The factor analysis suggests that there is a difference between the output and uptake of arguments, since level and content of justification load on one dimension, while respect and reciprocity load on another. Hence, the analysis of equality does not make use of an index with all four variables. Following Gerber (2015), a differentiation was made between the quality of contributions (level and content of justification) and considerations (respect and reciprocity).

Measuring deliberative equality

The capacity to engage in deliberation is associated with different characteristics, which individuals may possess to varying extents. However, due to the limited number of participants in the data set, it is necessary to refrain from including too many variables. Previous studies indicate that socio-demographic variables are important predictors for both discussion activity and discourse quality in citizen forums. In her study of participation in the Europolis deliberative opinion poll, Gerber (2015) finds that gender is an important predictor for how actively

participants engage in the discussions. Similar findings suggesting that women are less prone to speak up in group discussions have been presented in a number of studies (see Hastie, Penrod, & Pennington, 2013; Karpowitz et al., 2012). Caluwaerts (2012) and Han et al. (2015) also find that age, gender and education make a difference regarding the individual deliberative capacity of the participants in the forums they analyze.

Moreover, if we are interested in how people make reasoned arguments, we can hardly avoid discussing their cognitive predispositions. Hence, in addition to the basic socio-demographic variables (age, gender and education), measures of knowledge and political interest are included. According to Mendelberg (2002), people also vary in their need for cognition, i.e. their motivation to think in depth about the essential merits of a message. Consequently, people with a high need for cognition tend to generate more arguments. In order to discover whether having a well-defined opinion on the issue affects one's capacity for deliberation, a variable to measure attitude coherence has been included. This variable is based on eight items relating to nuclear power (from the pre-discussion survey), and coherence indicates that opinions with regard to nuclear power before the discussion started were situated predominantly in the pro or con camp.

Apart from their individual characteristics, the participants' capacity to take part in the deliberative process may also be affected by group-related factors. The experimental component in the deliberative forum from which the data originates was two different decision-making conditions, the vote and common statement. Even if the part of the deliberative process analyzed here only concerns the area where the small groups from both treatments had a parallel design (general discussion on nuclear power and its alternatives), the participants in the common statement groups knew that they would eventually have to come to some form of agreement. This naturally might have an effect on how they engage with other participants. As such, a treatment variable was included to account for contextual variation, and robust clustered standard errors were used in the regression analyses to account for variation at the group level. More information on the independent variables and how they are operationalized can be found in Appendix B.

Empirical Analysis

The analysis is divided into two parts. First, the results for the four basic elements of the deliberative process identified above are reported. Then the findings from the Finnish data are compared to the data from the three other deliberative forums and with data from the Steiner et al. (2004) findings on discourse quality in parliamentary debates. In the second part of the analysis, a more systematic analysis

of the participatory equality in citizen deliberation is presented. Here I rely on regression analyses in order to examine the equality of contributions and considerations in the deliberative process. Furthermore, the findings from three deliberative citizen forums are also compared to determine whether discussion activity and discourse quality can be explained the same way across different citizen forums.

The data from the Finnish deliberative forum suggest that most speeches with a demand included at least some attempt at justifying the demand to the other participants. Only 12 percent of the speeches with a demand involved no form of justification. Furthermore, 46 percent of the demands were supported by inferior justifications, and in 42 percent of the cases there was at least one qualified justification. These numbers bear a close resemblance to those that Caluwaerts (2012) reports for the Belgian deliberative forum; they are somewhat lower than those reported in the Europolis deliberative opinion poll (Steiner, 2012) and somewhat higher than those reported from the deliberative forums in Kansas (Han et al., 2015). Since comparisons across several categories would be almost impossible due to the varying number of categories used in each study, all comparisons are consequently made using only a category that can be matched across all studies.

Table 1. Distribution of DQI indicators and comparisons between deliberative citizen forums

INDICATOR	Finnish Deliberative Experiment	Europolis Deliberative Poll	Belgian Deliberative Forum	Kansas Deliberative Forum	Debates in parliament public/ non-p
Level of Justification	%	%	%	%	%
No Justification	12	–	–	–	–
Inferior Justification	46	–	–	–	–
Qualified (or better) Justification	42	60	41	29	88/60
Content of Justification					
Self or in-group interest	3	–	–	–	–
Neutral	90	–	–	–	–
Common good	7	18	5	9	31/9
Respect					
Explicit lack of respect	6	6	4	n/a	42/15
Respect (at least implicit)	94	–	–	–	–
Reciprocity					
Demand based on individual opinion	33	–	–	–	–
Demand with link to others opinions	57	–	–	–	–
Demand considers counter-argument	10	–	–	–	–

By comparing the numbers from the deliberative citizen forums to the study by Steiner et al. (2004) using DQI in a parliamentary setting, it was found that 88 percent of the opinions presented in plenary debates were supported by a complete justification. That is two times more than in the Finnish deliberative experiment and also a lot more than in the Europolis deliberative opinion poll. However, the differences are noticeably smaller if citizen deliberation is compared to the non-public debate in parliamentary committees, which is perhaps a better comparison to small-group discussions. In committees, the members of parliament make reason-based claims in 60 percent of the speeches (Steiner et al., 2004), which is the same as for the Europolis deliberative opinion poll (Steiner, 2012; Caluwarts, 2011).

According to deliberative theory, arguments should also be expressed in terms of the common good. However, when it comes to citizen deliberation, it appears that pleas to the common good are quite infrequent. In the Finnish case, the citizens in the small groups only made an appeal to the common good in about 7 percent of all arguments. This makes references to the common good about as (in)frequent as they were in the citizen deliberation on broadband access arranged in Kansas and in the Belgian experiment. The Europolis deliberative opinion poll differs somewhat from the other citizen forums with 18 percent of the arguments being expressed in terms of the common good. Nevertheless, the most substantial difference is the one between citizens and elected representatives. According to the study by Steiner et al. (2004), the concern for the common good during plenary debates in parliaments is over 30 percent. In committees, the quality of deliberation is again more similar to that of the citizen forums.

The data from the Finnish deliberative experiment suggest that there are a number of occasions in each group where participants are clearly disrespectful, although, in general, the level of respect is quite high. Only 6 percent of the arguments include an explicitly disrespectful remark. The discussions in both the Belgian deliberative forum and the Europolis deliberative opinion poll were also very respectful; only 5 percent of the speeches included a disrespectful remark (Steiner, 2012; Caluwaerts, 2012). However, in the plenary sessions in parliaments, disrespectful speeches were very common (42 percent), and while elected representatives again were more similar to ordinary citizens when debating in committees, they were still more likely to be disrespectful (15 percent). The last indicator, reciprocity — used to measure discourse quality in the Finnish deliberative forum — is not strictly comparable to any of the measures used by the other studies. According to this indicator, two-thirds of the demands are linked to a previous demand in one way or another.

However, only 10 percent of these demands actually engage with a counter-argument or make a comparison of different demands.

An average discourse quality measured for the whole forum does not tell us very much about the equality in the deliberative process. Hence, we need, beside the presence of different elements in deliberation, to know how these elements are distributed among the participants in the deliberative process. In the Finnish deliberative forum, the most active member in any of the small groups produced more than 50 arguments within the time frame for this analysis, while five participants did not produce a single argument. Although it is fairly obvious that some participants were quite dominating, the question remains as to how active someone should be in order to have any influence in the deliberative process at all. Should they produce one, five or ten arguments? Since it is very difficult to make any definite judgments on what levels of activity and discourse quality equal good deliberation, it makes more sense to see if there are any systematic differences in talk activity and discourse quality among the participants. In order to examine differences at the individual level, I make use of three regression models. The first model looks at what explains individual activity by analyzing the number of arguments produced by the participants. The second and third models are designed to explain individual variations in discourse quality.

In the first model (Table 2) a clear pattern can be discerned, suggesting that men and those with a higher education have dominated the discussion. Men and those with a higher education produce about 50 percent more arguments than women and those with the lowest education. Political interest and attitude coherence are also significant denominators for how actively participants put forward arguments in the deliberative process. The effect of the latter appears, however, to be rather small. Judging by the number of arguments presented, the deliberative process seems to suffer from a degree of inequality. Even though we cannot possibly expect everyone to act the same way or to be just as active in the deliberative process, the systematic differences in discussion activity suggest that all participants do not partake in the deliberative process on equal terms.

Table 2. Regression models (OLS) analyzing equality in deliberative process

	Number of arguments	Discourse Quality - Contributions	Discourse quality - Considerations
	B (standard error)	B (standard error)	B (standard error)
Age/100	0.91 (0.63)	-0.23** (0.07)	-0.15* (0.06)
Gender (male)	0.46*** (0.12)	0.00 (0.02)	-0.01 (0.03)
Education	0.62** (0.21)	0.01 (0.01)	0.03 (0.03)
Issue knowledge	0.28 (0.43)	0.10 (0.06)	0.07 (0.07)
Political interest	0.79** (0.24)	0.00 (0.05)	-0.05 (0.06)
Attitude coherence	0.04* (0.02)	0.05** (0.02)	-0.01 (0.03)
Treatment (common statem.)	0.13 (0.14)	-0.01 (0.02)	-0.01 (0.02)
Constant	0.49 (0.35)	0.61*** (0.05)	0.74*** (0.04)
R ² (adj.)	0.29	0.31 (0.25)	0.17 (0.10)
N	82	82	82
Note: All independent variables coded to vary between 0-1, number of arguments log transformed			
Levels of significance: *<0.1, **<0.05, ***<0.01			
Standard errors clustered at level of the small group			

Moving onto the discourse quality measures, the results are quite different from those of discussion activity. First of all, the explanatory variables do not indicate the same kind of systematic differences for discourse quality as they did for discussion activity. That being said, the model remains about as good at explaining discourse quality as it was at explaining the number of arguments, which would suggest that differences are much smaller for the quality than for the quantity of the

arguments. Nonetheless, some differences are to be found. It seems that younger people are better at generating high-quality arguments and more willing to consider other people’s viewpoints and arguments. Higher-quality contributions are also related to the participants’ attitude coherence. This finding follows the expectation that participants with a more coherent set of opinions and a stronger belief in their political abilities are more likely to produce reasons in support of their arguments (Mendelberg, 2002). While the results do not suggest that the participants with made-up minds are any less respectful or reciprocal, their deliberative capability seems to be limited to producing arguments for opinions they held at the outset.

The quality of the considerations is not related to any of the independent variables, apart from age. All in all, the explanatory power of the predictor variables is very much restricted to the output side of deliberation, be it discussion activity or discourse quality. If we compare the Finnish deliberative experiment to the other citizen forums, there are again many similarities. The socio-demographic variables explain both how likely people are to contribute and how likely they are to consider different proposals.

Table 3. Significant differences in deliberative equality for the different experiments

Finnish Deliberative Experiment	Europolis Deliberative Poll	Belgian Deliberative Forum	Kansas Deliberative Forum
(-) Age (-) Female (+) Education (+) Political interest	(-) Female (+) Political interest	(-) Age (-) Female (+) Education	(-) Age (-) Female (+) Education

Note: Includes only variables reported in more than case
-/+ = direction of relationship

The methods used to calculate individual levels of deliberative quality differs somewhat from case to case, making it hard to say exactly how great an effect each predictor had in a comparative sense. Nonetheless, it is quite clear that women and individuals with less education are at a disadvantage in the deliberative process. Based on the findings from the Europolis (Gerber, 2013) and the Finnish experiment, this disadvantage seems to be mainly related to the quantity of arguments. They simply produced fewer arguments, not lower-quality arguments per se. The findings from Caluwaerts (2012) and Han et al. (2015) indicate that these groups also produce somewhat lower discourse quality, but, as they do not

report the level of discourse quality in relation to activity, it is hard to judge whether their findings actually differ from the two other forums. Interestingly, younger participants show higher levels of discourse quality in three of the forums, and, perhaps less surprisingly, political interest predicts how actively people engage in the discussions.

Discussion

The aim of this study has been to examine the extent to which citizen deliberation fulfills the characteristics described in deliberative theory and whether the capacity for deliberative participation is equally distributed. Rather than examining only one specific deliberative mini-public, the findings from a Finnish citizen deliberation experiment have been compared with those from three other citizen forums where similar measurements have been used. The purpose of these comparisons was to increase the generalizability of findings on the deliberative capacity of ordinary citizens.

While there are some moderate differences between the different citizen forums, all four citizen forums display relatively moderate levels of justification, at least in comparison with elected representatives. On the other hand, the level of respect is much higher for all of the citizen forums than it is in debates involving elected representatives. It appears that the quality of justifications is at its highest and the level of respect is at its lowest in public forums, such as plenary discussions in parliament. Conversely, more informal forums — such as the Finnish experiment together with the deliberative forums in Belgium and the US — have poorer justifications, fewer references to the common good, but the highest levels of respect. This is perhaps not that surprising considering the different characteristics of the two fora. Politicians represent parties and constituents, and they cannot waver in their stance when they speak in public. Citizens discussing in small groups, on the other hand, benefit from listening and learning from each other and have much less to lose from adopting a new or different idea. Politicians are also more likely to reiterate prepared arguments, while citizens do not have the same level of experience in presenting political arguments. It also seems that references to the common good are more common when there is reason to believe that people might not have the common good in mind when they present their arguments. The citizen forums in the US, Finland and Belgium have substantially fewer references to the common good than parliamentary debates.

There are, however, also some more subtle differences in the data that are of interest. First, the differences between citizens and elected politicians were smaller when the representatives met in a non-public forum, such as a committee. Second,

Europolis deviated somewhat from the other citizen forums in terms of quality and references to the common good. In fact, the deliberative quality of the Europolis opinion poll was actually relatively close to that of elected representatives in committees. The reason why Europolis was a little different from the other citizen forums could be explained by the fact that it gathered participants from different countries, resulting in a greater variation in culture and languages that needed to be bridged. Hence, justifications and references to the common good may have had to be made more explicit than in the other citizen forums. Nonetheless, the distinctive difference is the one between citizen forums and plenary debates.

The analysis of deliberative equality suggested that women and individuals with less education are less likely to speak up during the deliberative process, while younger participants produce higher discourse quality than older participants. Other findings were somewhat mixed. In two of the cases, the differences in deliberative equality appeared to affect only the quantitative measure (discussion activity) and not the qualitative measure (DQI). The other two cases suggest that the differences might also apply to the discourse quality.

Comparing shares of high-quality arguments and respect in different forums is not without its problems, since it is difficult to estimate whether the differences among the forums are the result of differing contexts or merely differences in discussion dynamics. In a longer or more fast-paced discussion, there will be more speech acts and subsequently a lower share of high-quality input, even if the same number of arguments has been presented. Furthermore, the large discrepancy between the indicators of the DQI in a single forum is an indication that the elements of deliberation do not go together as well as normative theory would suggest. It is understandable that real-world deliberations will not live up to the ideals of deliberative democracy, at least for extended periods of time. However, this finding underlines the importance of looking at several dimensions of quality, before judging the success of a deliberative process. Merely looking at level of justification would suggest that parliamentary forums are more deliberative than citizen forums, while the level of respect gives the opposite impression.

Finally, a few words on the practical implications of these findings. As stated in the introduction of the article, deliberative practices have become increasingly popular in recent years. In this context it is comforting to know that well-designed deliberative forums appear to produce relatively similar deliberative quality in different contexts. This can help practitioners gain an understanding of what they need to do to achieve the type of productive deliberations they are looking for and what results they might expect from deliberative citizen forums. Based on the findings from this study, it seems as if achieving respectful discussions is less of a

challenge than upholding high levels of justification or giving everyone equal opportunity to make themselves heard.

The fact that deliberative citizen forums and parliamentary settings seem to capture different qualities of deliberation is also interesting with regard to the current debate on deliberative systems (Parkinson & Mansbridge, 2012). According to the deliberative systems approach, every part of a political decision-making process need not be perfectly deliberative for the process to bring about deliberative ends. The virtues of deliberation can be dispersed throughout the system and still contribute to a more deliberative democracy overall. From this perspective, it would make sense to find ways in to involve both citizen forums and elected bodies in political decision-making.

References

- Anckar, C. (2008). On the applicability of the most similar systems design and the most different systems design in comparative research. *International Journal of Social Research Methodology*, 11(5), 389-401.
- Andersen, V. N., & Hansen, K. M. (2007). How deliberation makes better citizens: The Danish deliberative poll on the Euro. *European Journal of Political Research*, 46, 531–556.
- Bohman, J. (1996). *Public deliberation: Pluralism, complexity, and democracy*. Cambridge, MA: MIT Press.
- Burkhalter, S., Gastil, J., & Kelshaw, T. (2002). A conceptual definition and theoretical model of public deliberation in small face-to-face groups. *Communication Theory*, 12(4), 398–422.
- Caluwaerts, D. (2012). *Confrontation and communication: Deliberative democracy in divided Belgium*. Brussels, Belgium: Peter Lang.
- Carson, L. (2006). Improving public deliberative practice: A comparative analysis of two Italian citizens' jury projects in 2006. *Journal of Public Deliberation*, 2(1).
- Chambers, S. (1996). *Reasonable democracy: Jürgen Habermas and the politics of discourse*. Ithaca, NY: Cornell University Press.
- Cohen, J. (1989). Deliberation and democratic legitimacy. In D. Matravers, & J. E. Pike (Eds.), *Debates in contemporary political philosophy: An anthology*. London: Routledge.
- De Vries, R., Stanczyk, A., Wall, I. F., Uhlmann, R., Damschroder, L. J., & Kim, S. Y. (2010). Assessing the quality of democratic deliberation: A case study of public deliberation on the ethics of surrogate consent for research. *Social Science & Medicine*, 70(12), 1896–903.
- Dryzek, J. S. (1990). *Discursive democracy: Politics, policy and political science*. Cambridge, England: Cambridge University Press.
- Dryzek, J. S. (2002). *Deliberative democracy and beyond: Liberals, critics, contestations*. Oxford, England: Oxford University Press.

- Dryzek, J. S., & Niemeyer, S. (2008). Discursive representation. *American Political Science Review*, 102(4), 481-493.
- Dutwin, D. (2003). The character of deliberation: Equality, argument, and the formation of public opinion. *International Journal of Public Opinion Research*, 15(3), 239-264.
- Farrar, C., Green, D., Green, J. E., Nickerson, D. W., & Shewfelt, S. (2009). Does discussion group composition affect policy preferences? Results from three randomized experiments. *Political Psychology*, 30(4), 615-647.
- Fishkin, J. S. (1991). *Democracy and deliberation: New directions for democratic reform*. New Haven, CT: Yale University Press.
- Fishkin, J. S. (1997). *The voice of the people: Public opinion and democracy*. New Haven, CT: Yale University Press.
- Fishkin, J. S. (2009). *When the people speak: Deliberative democracy and public consultation*. Oxford, England: Oxford University Press.
- Fung, A. (2007). Minipublics: Deliberative designs and their consequences. In S. Rosenberg (Ed.), *Deliberation, participation and democracy: Can the people govern?* (pp. 159-183). New York: Palgrave MacMillan.
- Gastil, J. (1993). *Democracy in small groups*. Philadelphia: New Society Publishers.
- Gastil, J., & Dillard, J. P. (1999). Increasing political sophistication through public deliberation. *Political Communication*, 16(1), 3-23.
- Geissel, B., & Newton, K. (2012). *Evaluating democratic innovations: Curing the democratic malaise?* London: Routledge.
- Gerber, M. (2015). Equal partners in dialogue? Participation equality in a transnational deliberative poll (Europolis). *Political Studies*, 63, 110-130.
- Gerber, M. (2013) "Examining the gold standard": *Deliberative quality in the transnational deliberative poll Europolis*. Bern, Switzerland: University of Bern

- Gerber, M., Bachtiger, A., Fiket, I., Steenbergen, M., & Steiner, J. (2014). Deliberative and non-deliberative persuasion: Mechanisms of opinion formation in EuroPolis. *European Union Politics*, 15(3), 410–429.
- Gerring, J. (2009). The case study: What it is and what it does. In C. Boix, & S. C. Stokes (Eds.), *Oxford handbook of comparative politics*. Oxford, England: Oxford University Press.
- Grönlund, K., Setälä, M., & Herne, K. (2010). Deliberation and civic virtue: Lessons From a citizen deliberation experiment. *European Political Science Review*, 2(1), 95-117.
- Gutmann, A., & Thompson, D. F. (1996). *Democracy and disagreement*. Cambridge, MA: Harvard University Press.
- Habermas, J. (1984). *The theory of communicative action — Lifeworld and system: A critique of functionalist reason*. Boston, MA: Beacon Press.
- Han, S., Schenck-Hamlin, W., & Schenck-Hamlin, D. (2015). Inclusion, equality, and discourse quality in citizen deliberations on broadband. *Journal of Public Deliberation*, 11(1).
- Hansen, K. M., & Andersen, V. N. (2004). Deliberative democracy and the deliberative poll on the Euro. *Scandinavian Political Studies*, 27(3), 261-286.
- Hastie, R., Penrod, S., & Pennington, N. (2013). *Inside the jury*. New Jersey: The Lawbook Exchange, Ltd.
- Himmelroos, S., & Christensen, H. (2014). Deliberation and opinion change: Evidence from a deliberative mini-public in Finland. *Scandinavian Political Studies*, 37(1), 41–60.
- Holzinger, K. (2004). Bargaining through arguing: An empirical analysis based on speech act theory. *Political Communication*, 21(2), 195–222.
- Karlsson, M. (2010, July 2). *What does it take to make online deliberation happen? A comparative analysis of 28 online discussion forums*. Paper prepared for presentation at the IV International Conference on Online Deliberation, Leeds, England.

- Karpowitz, C. F., Mendelberg, T., & Shaker, L. (2012). Gender inequality in deliberative participation. *American Political Science Review*, 106(3), 533–547.
- Knight, J., & Johnson, J. (1997). What sort of equality does deliberative democracy require? In J. Bohman, & W. Rehg (Eds.), *Deliberative democracy: Essays on reason and politics*. Cambridge, MA: MIT Press.
- Kohn, M. (2000). Language, power, and persuasion: Toward a critique of deliberative democracy. *Constellations*, 7(3), 408–429.
- Luskin, R. C., Fishkin, J. S., & Jowell, R. (2002). Considered opinions: Deliberative polling in Britain. *British Journal of Political Science*, 32(3), 455–487.
- Mendelberg, T. (2002). The Deliberative citizen: Theory and evidence. *Political Decision Making, Deliberation and Participation*, 6, 151–193.
- Nabatchi, T. (2010). Deliberative democracy and citizenship : In search of the efficacy effect. *Journal of Public Deliberation*, 6(2).
- Parkinson, J., & Mansbridge, J. (2012). *Deliberative systems: Deliberative democracy at the large scale*. Cambridge, England: Cambridge University Press.
- Posner, R. A. (2003). *Law, pragmatism, and democracy*. Cambridge, MA: Harvard University Press.
- Przeworski, A. (2010). *Democracy and the limits of self-government*. Cambridge, England: Cambridge University Press.
- Ryan, M. (2014). *Advancing comparison of democratic innovations: A medium-N fuzzy-set qualitative comparative analysis of participatory budgeting*. Southampton, England: University of Southampton.
- Ryan, M., & Smith, G. (2014). Defining mini-publics. In *Deliberative mini-publics: Involving citizens in the democratic process* (pp. 9–26). Colchester, England: ECPR Press.
- Ryfe, D. M. (2005). Does deliberative democracy work? *Annual Review of Political Science*, 8(1), 49–71.

- Sanders, L. M. (1997). Against deliberation. *Political Theory*, 347, 1–17.
- Setälä, M., Grönlund, K., & Herne, K. (2010). Citizen deliberation on nuclear power: A comparison of two decision-making methods. *Political Studies*, 58, 688–714.
- Smith, G. (2009). *Democratic innovations: Designing institutions for citizen participation*. Cambridge, England: Cambridge University Press.
- Steenbergen, M. R., Bächtiger, A., Spörndli, M., & Steiner, J. (2003). Measuring political deliberation: A discourse quality index. *Comparative European Politics*.
- Steiner, J. (2012). *The foundations of deliberative democracy: Empirical research and normative implications*. Cambridge, England: Cambridge University Press.
- Steiner, J., Bächtiger, A., Spörndli, M., & Steenbergen, M. R. (2004). *Deliberative politics in action: Analyzing parliamentary discourse*. Cambridge, England: Cambridge University Press.
- Stromer-Galley, J. (2007). Measuring deliberation's content : A coding scheme. *Journal of Public Deliberation*, 3(1).
- Young, I. M. (2002). *Inclusion and democracy*. Oxford, England: Oxford University Press.

APPENDIX A – Discourse Quality Index (DQI)

Indicator	RCA*
<i>I. Level of Justification</i>	
0: no justification, participant presents only his/her point of view 1: inferior justification; conclusion(s) embedded in (an) incomplete inference(s), no linkage is made as to why X will contribute to Y 2: qualified justification; one (or more) conclusion embedded in a complete inference, a linkage is made as to why X will contribute to Y (other incomplete inferences may be present)	0,75
<i>II. Content of Justification</i>	0,93
0: explicit statement concerning group/self- interest 1: neutral statement; no reference to group or self-interest, but no reference to common good either 2a: explicit reference to common good in utilitarian or collective terms 2b: explicit statement in terms of the common good with reference to the difference principle	
<i>III. Respect</i>	0,96
0: disrespect; explicitly negative statement concerning either the group/person under discussion or the other participants and their views 1: implicit respect; no explicitly negative statement concerning the group/person under discussion nor the other participants and their views	
<i>IV. Reciprocity</i>	0,82
0: No connection to argument/demand presented by another participant 1: Connects directly to an argument presented by another participant 2: Considers a counter-argument in own argumentation or compares/weighs different arguments	

*RCA – Ratio of Coder Agreement. Intersubjective reliability of the coding was tested by looking at 84 speech acts from four randomly selected segments.

APPENDIX B – Variables and descriptive statistics

Variable	Coding	Descriptive information					
		<i>N</i>	<i>Min</i>	<i>Max</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>VIF</i>
Speech activity	Number of arguments produced by each participant in the deliberative experiment. Log transformed in regression analysis due to skewed distribution.	88	0	51	13.2	10.2	
Deliberative output	Additive index combining two aspects of DQI: I. Level of Justification II. Common good. Vary between 0 and 1 (highest quality).	85	0.33	0.88	0.59	0.09	
Deliberative uptake	Additive index combining two aspects of DQI: III. Respect, IV. Reflection. Vary between 0 and 1 (highest quality).	85	0.45	0.88	0.67	0.08	
Age	Age in years divided by 100.	90	0.18	0.72	0.48	0.16	1.09
Gender	0: female, 1 = male.	90	0.00	1.00	0.53	0.50	1.40
Education	Highest level of schooling completed; 0:	90	0.00	1.00	0.47	0.31	1.19

	Basic schooling, 1: Secondary, 2: Lower tertiary, 3: Higher tertiary. Recoded to vary between 0 and 1 (highest).						
Political interest	Item from pre-deliberation survey: <i>How interested are you in politics?</i> 5-point scale varying between <i>Very interested</i> and <i>Not at all interested</i> .	88	0	1	0.62	0.23	1.11
Energy knowledge	Number of correct answers to 10 factual questions on nuclear energy use; coded to vary between 0-1 (1= all correct).	90	0.00	0.90	0.53	0.18	1.56
Attitude coherence	Measures whether the respondent has a coherent attitude towards the eight nuclear power items. When all statements indicate either a positive or a negative attitude towards	89	0.00	1.00	0.64	0.35	1.10

	nuclear power, the participant has a strongly coherent opinion. Varies between 0-1 (1 = high coherence).						
Common statement	Whether group decision to be made with common statement or vote (1=common statement).	90	0.00	1.00	0.50	0.50	1.07

APPENDIX C - Factor analysis of *DQI* indicators

DQI indicators	Components	
<i>I. Level of Justification</i>	0.794	
<i>II. Content of Justification</i>	0.738	
<i>III. Respect</i>		0.608
<i>IV. Reciprocity</i>		0.825
Eigenvalue	1.25	1.03
% Variance explained	30.52	26.52

Extraction Method: Principal Component Analysis with Varimax rotation.