

Deliberative Abilities and Influence in a Transnational Deliberative Poll (EuroPolis)

MARLÈNE GERBER, ANDRÉ BÄCHTIGER, SUSUMU SHIKANO,
SIMON REBER AND SAMUEL ROHR*

This article investigates the deliberative abilities of ordinary citizens in the context of ‘EuroPolis’, a transnational deliberative poll. Drawing upon a philosophically grounded instrument, an updated version of the Discourse Quality Index (DQI), it explores how capable European citizens are of meeting deliberative ideals; whether socio-economic, cultural and psychological biases affect the ability to deliberate; and whether opinion change results from the exchange of arguments. On the positive side, EuroPolis shows that the ideal deliberator scoring high on all deliberative standards does actually exist, and that participants change their opinions more often when rational justification is used in the discussions. On the negative side, deliberative abilities are unequally distributed: in particular, working-class members are less likely to contribute to a high standard of deliberation.

Keywords: deliberation; deliberative polls; European Union; citizen participation; working class.

Confronted with a malaise of democratic governance and the disenchantment of citizens with politics, recent years have witnessed a worldwide boom in participatory and deliberative citizen events. The idea is that democratic innovations might not only help to narrow the gap between politicians and citizens but also serve as a policy-consultation device. Yet, if democratic innovations should become a regular component of democratic governance, it is essential to know whether they do in fact function as their proponents suggest. One crucial question in this regard is whether ordinary citizens can deliberate together at high quality levels. For James Fishkin, one of the deliberative pioneers, the lessons from the manifold experiences with citizen deliberation worldwide are clear: ‘everybody can deliberate’.¹ Fishkin is not the only one claiming that ordinary

* Institute of Political Science, University of Bern; Institute of Social Sciences, University of Stuttgart; Department of Politics and Public Administration, University of Konstanz; Institute of Political Science, University of Bern; The Graduate Institute, University of Geneva (emails: marlene.gerber@ipw.unibe.ch; andre.baechtiger@sowi.uni-stuttgart.de; susumu.shikano@uni-konstanz.de; sim.reber@gmail.com; samuel.rohr@graduateinstitute.ch). The authors would like to thank participants in the panel ‘Deliberation and Democratic Legitimacy’ at the MPSA Conference 2011, participants at the workshop ‘Frontiers of Deliberation’ at the ECPR Joint Sessions of Workshops in St. Gallen 2011 and the panel ‘Committees and Rational Decision-Making’ at the ECPR General Conference in Reykjavik 2011, participants at seminars at the University of Frankfurt and Oldenburg (2012), three anonymous reviewers, and the editors Bob Goodin and Hugh Ward for excellent comments and suggestions on previous versions of this article. The authors are particularly grateful to John Dryzek and Jane Mansbridge for their extremely valuable feedback, their encouragement and their support, and also wish to thank David Sanders for helping us replicate his findings on opinion change in EuroPolis. Finally, they would also like to thank Karolina Kojder for her careful work in transcribing and translating the Polish group discussions. This project was funded by the Swiss National Science Foundation (‘Potential for Deliberation among EU Citizens’, no. 100012_126483). André Bächtiger and Susumu Shikano also thank the Hanse Wissenschaftskolleg (Delmenhorst, Germany) for its support during this project. Data replication sets are available at <https://dataverse.harvard.edu/dataverse/BJPolS>. Online appendices are available at <http://dx.doi.org/doi:10.1017/S0007123416000144>.

¹ University of Texas (2003). An Expertise in Democracy. At <http://www.utexas.edu/features/archive/2003/polling.html>, retrieved February 27, 2013.

citizens can be turned into good deliberators. The thrust behind the deliberative movement is that Schumpeterian conceptions of the minimalist democratic citizen are woefully wrong and that, with a little helping hand, ordinary citizens can approach the philosophical ideals of deliberation.²

But the deliberative gospel has not convinced everyone. Since the advent of deliberative theories in the 1990s, deliberation has met with sustained criticism. Critiques have revolved around deliberation's inconsistency with psychological theories and experiences of human action.³ Critics have, first of all, questioned the idealized conjectures of the 'deliberative citizen' possessing sophisticated reasoning skills while being simultaneously respectful, reflective, inquisitive and open-minded. Drawing upon the experience of psychological experiments and jury deliberations, critics also claim that deliberative abilities are strongly correlated with socio-economic, cultural and psychological factors.⁴ If such distortions in deliberative ability exist and the outcomes of deliberation are simultaneously driven by non-deliberative pathways, then the deliberative ideal of an egalitarian and unconstrained exchange of arguments is violated. In this case, citizen deliberation would boil down to an undemocratic exercise, giving deliberative advocates a hard time claiming relevance and legitimacy for deliberation's outcomes and especially for using them as policy-consultation devices.

So far, however, a great deal of this controversy is surprisingly theoretical rather than being based on systematic empirical facts. Many critics tend to overlook the fact that deliberation today is a highly structured affair. Most citizen deliberations are conducted under *supportive conditions*, i.e., citizens get balanced information material, experts answer citizens' questions, and facilitators ensure that small group discussions keep to the topic and are focused on all the arguments. Thus, drawing far-reaching implications from psychological experiments and jury deliberations that lack such supportive underpinnings may be defective. Conversely, deliberationists tend to assume that reasoned deliberation will quasi-automatically follow when conditions for deliberation are good. This may explain why 'researchers have been less interested in deliberation itself than in measuring its effects'.⁵ But they have downplayed the possibility that, even under optimal conditions, deliberation may not occur as expected by normative theory. Not only may some citizens be overburdened by the deliberative process, citizen 'deliberation' may also be just talk, in the absence of any philosophical underpinnings. To date, existing studies portray a fairly positive image of the quality of citizen deliberation under supportive conditions; yet analyses are generally based on self-perceptions of participants which may be fraught with 'social desirability' issues.⁶ Existing studies analysing deliberative quality on an external basis find mixed results.⁷ However, these studies have only focused on a limited range of indicators of deliberative quality, and none of them have asked the question whether and to what extent citizens actually possess the required abilities to reach various deliberative virtues. What is more, there are very few studies analysing the deliberative process in deliberative polls (DPs), which some have called the 'gold standard' among contemporary mini-publics⁸ developed by James Fishkin and his collaborators.⁹

² See, e.g., Dryzek 2010.

³ E.g., Mutz 2008.

⁴ See, e.g., Mendelberg 2002.

⁵ Ryfe (2005), p. 54.

⁶ Andersen and Hansen 2007; Fishkin, Luskin and Siu 2014; Grönlund, Setälä and Herne 2010; Ratner 2008; see also Sanders 2012.

⁷ E.g., Dutwin 2003; Himmelroos and Christensen 2014; Pedrini 2014; Rosenberg 2007; Schneiderhan and Khan 2008; Siu 2009; Stromer-Galley 2007.

⁸ Mansbridge 2010.

⁹ Fishkin 1995; see, e.g., Andersen and Hansen 2007; Fishkin 2009; Fishkin et al. 2010; Luskin, Fishkin and Jowell 2002; Luskin et al. 2014.

Surely, from a Habermasian viewpoint, deliberative mini-publics may not form any 'gold standard' for good citizen deliberation; rather, the Habermasian vision has a critical theory angle, where contestation and emancipation in the wider public sphere are central.¹⁰ Nonetheless, deliberative mini-publics and DPs in particular have evolved as one key model of how to organize citizen deliberation in practice and its institutional precepts have been replicated on a world-wide basis. While we know that the DP fares well on the input-criteria of representation and on several output-criteria, such as opinion change, knowledge gain and satisfaction with participation in the event, we know surprisingly little about how participants deliberate during a DP event.¹¹

Our article looks inside the black box of deliberation of the DP and asks whether the ideal deliberator, scoring high on important deliberative standards, actually exists when 'insulated from certain negative or distorting effects of the broader public sphere'.¹² In concrete, we focus on 'EuroPolis', a pan-European deliberative poll, which was carried out in Brussels in May 2009.¹³ To measure deliberative quality, we utilize an updated version of the Discourse Quality Index (DQI) which employs a broad understanding of deliberative quality and allows for a quantitative content analysis of recorded discussions.¹⁴ The aggregation of the different DQI components is accomplished via (Bayesian) item response analysis.¹⁵ The Item Response Theory (IRT) provides a novel way of conceptualizing deliberative quality, by exploring how well citizens are able to achieve the various standards of deliberative quality (justification rationality, common good orientation, or respect) and whether the standards form a conceptual whole, while simultaneously making the realistic assumption that some components of deliberative quality may be more difficult to achieve than others. By mapping all participants as ideal points on latent dimensions of deliberative quality, IRT not only enables us to check whether 'deliberative citizens'¹⁶ exist, but also to analyse whether deliberative ability is associated with socio-economic, cultural and psychological factors. Moreover, we also investigate pathways of deliberative influence, exploring whether argument-based factors rather than non-deliberative dynamics and other distortions drive opinion change.¹⁷

The remainder of the article is organized as follows. The next section gives more background on the controversy between critics and advocates of deliberation. There follows the presentation of an updated version of the Discourse Quality Index, defining thresholds for high quality deliberation, and introducing the aggregation method, and Bayesian item response analysis. The next section gives some background on EuroPolis, and describes the data. The empirical findings and the conclusion follow.

¹⁰ We thank an anonymous reviewer for highlighting this crucial aspect.

¹¹ See Siu (2009) for an exception of analysing online deliberative polls. However, Siu considers the indicator of justification-rationality, but not of respect or common-good orientation, for instance.

¹² Chambers (2004), p. 400.

¹³ EuroPolis is a special case of deliberative polling (but see also Fishkin (2009) for the first transnational deliberative poll 'Tomorrow's Europe'), since it involves participants from many different nationalities, requiring that the discussions were simultaneously translated. While the highly diverse nature of the EuroPolis event might imply more difficult conditions for citizen deliberation than a classic deliberative poll, Doerr (2009) shows that simultaneous translation can promote both listening and inclusiveness.

¹⁴ Steiner et al. 2004.

¹⁵ Clinton, Jackman and Rivers 2004; Jackman 2001; Rasch 1980; Shikano 2008.

¹⁶ Mendelberg 2002.

¹⁷ When we speak of deliberative ability, we do not see this as a fixed and unchangeable trait, but rather as a *communicative repertoire* at which some people may be better than others, but that every individual can eventually cultivate and develop.

THE DELIBERATIVE INCOMPETENCE AND DISTORTION THESIS

From a classic perspective, ideal deliberators must fulfil demanding behavioural obligations: they must be reasoned, common-good orientated, reflective, respectful, empathetic, inquisitive and open to the better argument. Many psychologists and those sceptical of deliberation have argued that only a small minority of individuals possesses the level of deliberative ability required by classic deliberative theory and that these abilities are also correlated with socio-economic and other factors.¹⁸ Since a key goal of deliberation is to include ‘all affected interests’ and empower the disenfranchised, a tension between inclusion and deliberative ability may arise, turning deliberation into a potentially harmful intervention that further marginalizes already disenfranchised groups. We summarize the various challenges under the label of ‘deliberative incompetence and distortions thesis’.¹⁹ This thesis has several dimensions, ranging from the general deliberative abilities of ordinary citizens and the ‘unitary deliberator model’ to social, cultural and psychological distortions, and to distortions in the outcomes of deliberation (which we address separately in the section on deliberative influence). We will now address these various aspects of the ‘deliberative incompetence and distortion thesis’ in turn.

General Deliberative Abilities and the ‘Unitary Deliberator Model’

Drawing on social and cognitive psychology, Rosenberg argues that most citizens lack the general abilities to participate in high-quality deliberation: ‘most “participants” who attend a deliberation do not, in fact, engage in the give and take of the discussion’. Rather, they ‘offer simple, short, unelaborated statements of their views of an event’.²⁰ Social psychology, so named by Rosenberg, suggests that this lack of abilities is not just ‘circumstantial’ and a matter of inadequate information or motivation, but that these limits are inherent and hard-wired. This is exactly the point of controversy between advocates and psychological critics of deliberation: while the former claim that supportive institutional devices – such as information provision – can help citizens to approach deliberative ideals, the latter deny this possibility. To date, however, we largely lack systematic empirical analysis of citizens’ deliberative abilities under *supportive conditions*.

Another question is whether a ‘unitary deliberator’²¹ simultaneously scoring high on all standards of high quality deliberation (justification rationality, respect, etc.) can exist in reality. Notice first that some deliberative theorists have argued that deliberation must not be conceived of as a ‘single evaluative whole’,²² since deliberative virtues may be unevenly distributed across various *macro-level sites* of the democratic system. This may also be true at the group level: as long as reasons are given, acknowledged and integrated into the discussion and recommendations, then it does not matter whether each individual possesses advanced deliberative abilities.²³ While the last part of this article explores deliberation’s role as a group resource, we are nonetheless interested in the existence of the ‘deliberative citizen’, a topic largely neglected in the literature. At the level of individuals, uni-dimensionality of deliberative quality may still matter: if the latter falls into its diverse components (technically speaking, is multi-dimensional), then we are likely to capture something other than a true ‘deliberative personality’. For instance, if actors only justify

¹⁸ Rosenberg 2014; Sanders 1997.

¹⁹ We borrow the term ‘deliberative distortions’ from Luskin et al. 2015.

²⁰ Rosenberg (2014), p. 108.

²¹ Goodin 2005.

²² Goodin 2005; Thompson 2008.

²³ We thank an anonymous reviewer for highlighting this point. See also Karpowitz and Raphael (2014), pp. 238–9.

their positions but never respond to other positions with respect, then this is an indication of rhetorical rather than true deliberative action.²⁴ Secondly, the existence of the ‘deliberative personality’ can also be questioned from the view of personality psychology.²⁵ Jennstal argues that an ‘ideal deliberator’ must score high on at least three personality traits, namely extraversion, agreeableness and openness; according to Jennstal, reason-giving falls primarily under the rubric of extraversion, whereas reflexivity, respect, empathy and inquisitiveness, in turn, require openness and/or agreeableness.²⁶ This, however, represents a rare constellation of personality traits: for instance, being extraverted does not necessarily imply that one also scores high on agreeableness. Certainly, personality psychology does not rule out the possibility of a ‘deliberative personality’ scoring high on all quality standards of deliberation; it just puts some serious question marks on its existence. Unfortunately, the EuroPolis questionnaire does not include personality questions; so we cannot directly link personality traits to deliberative behaviour. But we can indirectly explore whether true deliberators exist, by checking whether the different components of deliberative quality – justification rationality, respect, empathy and inquisitiveness – form a compound and uni-dimensional phenomenon at the level of deliberating citizens.

Socio-economic, cultural and psychological distortions. Scholars studying political behaviour have long demonstrated that socio-economic and cognitive *resources* influence an individual’s ability to get politically engaged.²⁷ Given the fact that deliberation is a more demanding form of participation than electoral participation,²⁸ many critics of deliberation expect that differences in enabling resources will play an even more important role in a deliberative event.

With regard to socio-economic factors, difference democrats and feminists have claimed that rational argumentation including logical deduction and general principles is frequently associated with men and socially privileged groups, while more tentative, figurative and emotional forms of expression are often associated with women, socially less privileged groups and cultural minorities.²⁹ This (bold) argument requires some qualifications. Regarding gender, there is an important counterargument. Some scholars have argued that women have higher capacities for respect and empathy and thus may actually be better deliberators than men.³⁰ Regarding class, sociological research has claimed that there may be class-specific ways of speaking and arguing. According to Bernstein, working-class people tend to adopt a restricted code of speech by using rather simple, repetitive and limited vocabulary that stands in contrast to the more accurate and elaborate code of speech employed by the middle class.³¹ Class differences in the ability to speak and argue are also closely linked to differences in education. On the one hand, well-educated people have access to occupations where they can develop reasoning and public-speaking skills.³² On the other hand, education may render people more ‘democratically enlightened’ in that they may display a higher adherence to democratic values and a better understanding of alternative preferences and positions.³³ Age, in turn, can be seen

²⁴ Jennstal and Niemeyer 2014.

²⁵ See, e.g., McCrae and Costa 1987.

²⁶ Jennstal (2012), pp. 17–18.

²⁷ E.g., Verba, Schlozman and Brady 1995.

²⁸ Hooghe 1999.

²⁹ Sanders (1997); Young (2002), pp. 38–40.

³⁰ Norris (1996), p. 91.

³¹ Bernstein 1971.

³² Mendelberg 2002.

³³ Nie, Junn and Stehlik-Barry (1996) demonstrate this for American citizens using data from the General Social Surveys and the National Election Studies (1972–94).

as a proxy for experience with political affairs. Experience may increase a person's ability for self-reflection and responsiveness to others.³⁴ Empirically, the role of socio-economic factors is mixed. Based on experimental studies, psychologists found that men and people from upper classes with higher education levels speak more frequently, stay more focused on the topic and contribute more varied and more relevant statements.³⁵ By contrast, the few scholars focusing on the process of citizen deliberation generally did not find massive socio-economic stratification of deliberative behaviour.³⁶

Regarding cultural distortions, theorists of multiculturalism worry that the deliberative ideal of rational argumentation represents a culturally specific format of communication.³⁷ In our sample of European citizens, 'culture' mainly concerns cultural differences between Southern, Central and Eastern, and Western Europeans. With regard to Southern Europeans, Gambetta has put forward a highly controversial argument, namely that Southern European societies feature 'Claro!' cultures in which admitting uncertainty or lack of knowledge is considered a weakness.³⁸ This, in turn, undermines deliberative ideals such as open-mindedness or respect. Following Gambetta's 'essentialist' argument, Southern Europeans should exhibit a lower quality of deliberation. With regard to Eastern Europeans, several studies indicate that citizens from Central and Eastern European countries have lower levels of 'republicanism' than citizens from Western European countries: they are less interested in politics, less engaged in civic affairs, and also have less trust in others.³⁹ This cultural dividing line may also translate into different deliberative behaviour, with the (tentative) expectation that Central and Eastern Europeans might perform less well than Western Europeans. Empirically, several researchers have identified traces of cultural differences in deliberative behaviour, finding that some societal cultures are less compatible with deliberative ideals.⁴⁰ Yet alleged cultural differences may sometimes merely reflect different experiences with deliberative practices. Focusing on elite deliberation in working groups of the European Council of Ministers, Naurin finds that new member states from Central and Eastern Europe displayed lower levels of deliberative quality than old member states. According to Naurin, this has to do with experience rather than culture, since old member states are more accustomed to playing 'the Brussels game'.⁴¹

Let us finally turn to *psychological distortions*. Given the limitations of the EuroPolis questionnaires, our study focuses only on motivation, involvement and knowledge. First, motivation and involvement are crucial factors for deliberative performance. High motivation and high involvement lead to 'central reasoning', leading to a willingness to diligently consider information and arguments; by contrast, low motivation and low involvement is conducive to 'peripheral reasoning' and reliance on information shortcuts.⁴² There is, however, some controversy whether high or low involvement increases or decreases deliberative quality. Fung acknowledges that one possibility is that individuals with low stakes in a discussion (albeit with a basic motivation to engage with the topic under discussion) will be the better deliberators, since low stakes are conducive to dispassionate attitudes and open-mindedness.⁴³ This view is

³⁴ Rosenberg 2005.

³⁵ Hastie, Penrod and Pennington 1983.

³⁶ Kim, Siu and Sood (2010), Pedrini (2014), Siu (2009); a partial exception is Hansen (2004).

³⁷ See, e.g., Young 2002.

³⁸ Gambetta 1998.

³⁹ Fuchs and Klingemann 2002; Karp and Banducci 2007.

⁴⁰ E.g. Duchesne and Haegel 2007; Min 2009.

⁴¹ Naurin (2010), p. 47.

⁴² Petty, Cacioppo and Schumann 1983.

⁴³ Fung 2003.

in line with classic deliberative theory, emphasizing calm and dispassionate reasoning. But Fung suggests that the opposite might also be true: participants with high stakes may ‘invest more of their psychic energy and resources into the process and so make it more thorough and creative’.⁴⁴ Secondly, knowledge about the topic may influence deliberative behaviour as well. Participants with a higher level of prior knowledge about the issue at hand may have a broader argumentative repertoire, which may positively influence their deliberative behavior.⁴⁵

DELIBERATIVE INFLUENCE

While our micro-level focus precludes us from judging the deliberative quality of the event as a whole, we nonetheless consider some key aspects of the outcomes of deliberation. In this article, we concentrate on one outcome that has been at the forefront of citizen deliberation and particularly of deliberative polling, namely opinion change. When opinion change takes place, deliberative theory would require it to occur ‘via mechanisms specified in the normative theories’.⁴⁶ In a Habermasian understanding of deliberation, only attempts to ‘convince each other that there are inherently good reasons to pursue one course of action over another’ justify a change in opinions.⁴⁷ Since it is exceedingly difficult to define what ‘good reasons’ are, we focus on well-justified arguments as a proxy variable.⁴⁸ By well-justified arguments, we expect that extended linkages are made between a premise and a conclusion, so that other participants can better judge the rationales behind a position.⁴⁹ We expect that well-justified arguments serve as *group resource* and affect opinion formation.

While the psychological literature does not exclude the possibility of systematic, argument-based opinion change, it also emphasizes non-deliberative pathways to opinion formation, such as undesired group dynamics where initial opinion distributions in the discussion group affect post-deliberative opinion. Overall, it is normatively questionable when participants’ post-deliberative opinions are not affected by argument-based pathways at all. Surely, from a democratic perspective, a problematic scenario arises when advantaged and rhetorically high-skilled participants *regularly* impose their pre-deliberative views on other participants, without being open to other participants’ viewpoints. The regularity criterion, however, cannot be tested in a single event but requires a meta-analysis of a large number of deliberative mini-publics. Moreover, as long as advantaged participants listen to others – a crucial condition of the ‘deliberative citizen’ specified above – and are open to changing their minds, it may not be problematic if their higher quality justifications serve as an epistemic group resource and affect opinion formation in other participants. In the following, we will explore the various dimensions of the ‘deliberative incompetence and distortion thesis’ empirically, by developing a new measure of deliberative ability and by focusing on the deliberative abilities of ordinary citizens and their consequences under supportive institutional conditions.

MEASURING DELIBERATIVE QUALITY

In the past decade, several scholars have explored the ‘deliberative incompetence and distortion thesis’ empirically. But they have been mostly concerned with the input dimension, i.e., the

⁴⁴ Fung (2003), p. 345.

⁴⁵ Cappella, Price and Nir 2002.

⁴⁶ Neblo (2007), p. 1.

⁴⁷ Chambers (1996), p. 99. See also Barabas 2004; Fishkin 2009; Mansbridge et al. (2012), p. 11.

⁴⁸ See Siu 2009.

⁴⁹ Chambers 1996; Landemore and Mercier 2012.

inclusiveness of the deliberative process, and the output dimension, i.e., the question whether deliberation leads to (unbiased) opinion changes.⁵⁰ While some studies have included measures of process quality, this is mostly done via survey-based self-reports of participants.⁵¹ But this approach is problematic as well: not only may self-reports contain elements of social desirability, they may also insufficiently capture the philosophical ramifications of the deliberative model. For instance, participants may think that the quality of reasoning was good, whereas philosophers would judge the respective reasoning as insufficient by their own theoretical standards. Only recently, researchers have begun to evaluate the quality of deliberation among ordinary citizens on the basis of an external and philosophically grounded measure.⁵² Compared to these pioneering attempts, our approach provides a more comprehensive measure of deliberative ability, by considering a larger batch of deliberative indicators while simultaneously setting a threshold for high and low deliberative quality and employing novel aggregation techniques.

Updated Discourse Quality Index

We assess deliberative quality on the basis of an updated version of the *Discourse Quality Index* (DQI).⁵³ The DQI allows for a quantitative content analysis at the level of individual speeches of recorded discussions. While there exist other measures of deliberative quality to study discussions among citizens,⁵⁴ we concentrate on the DQI, for two reasons. First, not only is there some convergence on what counts as high quality deliberation (such as reason-giving and reciprocity), the DQI has also met with considerable support from prominent deliberative philosophers.⁵⁵ However, the original DQI was developed for the analysis of parliamentary debates and is strongly rooted in a classic and Habermasian-inspired understanding of deliberation, emphasizing rational argumentation. This raises questions of how well the DQI can be applied to citizen deliberation. As our empirical analysis will demonstrate, such concerns are unsubstantiated.⁵⁶ Nonetheless, we make some adaptations in the evaluation procedure to take into account the constraints of citizen deliberation (see below). In this regard, the DQI can also profit from being enriched with ‘alternative’ forms of communication. Many scholars nowadays consider ‘alternative’ communication modes such as story-telling or testimony as fully valid and even desirable deliberative practices.⁵⁷ We think that such developments in deliberative theory and practice must be reflected in an evaluation of citizens’ deliberative abilities as well and therefore add one element of expanded notions of deliberation, namely ‘story-telling’. In the following, we briefly describe the various components of the updated Discourse Quality Index (for coding examples, see Appendix, Table A2).

Justification rationality. A core indicator of deliberative quality is reason-giving. Since the ideal speech situation itself has no content, one cannot apply external standards to what constitutes a good reason. Hence, we focus on the syntactic structure of argument and judge to what

⁵⁰ E.g., Barabas 2004; Luskin, Fishkin and Jowell 2002.

⁵¹ E.g., Andersen and Hansen 2007; Sanders 2012.

⁵² E.g., Himmelroos and Christensen 2014; Pedrini 2014; Siu 2009.

⁵³ Steiner et al. 2004.

⁵⁴ E.g., Andersen and Hansen 2007; Dutwin 2003; Rosenberg 2007; Stromer-Galley 2007.

⁵⁵ Habermas 2005; Thompson 2008.

⁵⁶ See also Siu (2009), Caluwaerts (2012), Steiner (2012), Himmelroos and Christensen (2014), Pedrini (2014) for successful application of (elements of) the DQI to citizen deliberation.

⁵⁷ See Bächtiger et al. 2010.

extent a speaker gives complete justifications and thus makes his speech accessible to rational critique. We distinguish among four levels of justification rationality: (0) no justification; (1) inferior justification where the linkage between reasons and conclusion is tenuous (this code also applies if a conclusion is merely supported with illustrations); (2) qualified justification where a linkage between reasons and conclusion is made; (3) sophisticated justifications where a problem is examined in-depth by providing various, well-justified arguments.

Common good orientation. Many deliberative democrats emphasize that arguments should be formulated with an eye on what we have in common and what is universal. We measure whether arguments are cast in terms of narrow group or constituency interests, whether there is neutral reference or mixed reference (i.e., reference to both narrow group interest and the common good), or whether there is a reference to the common good. In the context of a pan-European discussion such as EuroPolis, the categories need to be refined, however. We distinguish between references to country interests (coded 1), references to two sorts of interests, country and European interests (scored 1.5), European Union (coded 2) and world community interests (coded 3), and the absence of such references (coded 0).

Respect towards other participants' arguments. Good deliberation is not only about mutual reason-giving with a focus on the common good, it also implies listening and 'uptake' of others' arguments with respect. We measure whether speakers include other participants' arguments but degrade them (coded 0), whether speakers ignore other participants' arguments (coded 1), whether they include other arguments in a neutral fashion (coded 2), and whether they value other participants' arguments (coded 3).

Respect towards groups. Deliberative quality also entails that participants show empathy and 'take into account the goals or values of persons unlike themselves'.⁵⁸ In the context of the EuroPolis discussions on immigration, this concerned third-country migrants. We capture whether speakers denigrate migrants (scored 0), don't refer to them (scored 1), whether they make reference to migrants in a neutral fashion (scored 2) or whether they show explicit respect towards them (scored 3).

Questioning. Deliberative democrats also emphasize the importance of inquisitiveness.⁵⁹ We operationalize inquisitiveness via questioning. Questioning has an informational and a critical function, even though the two frequently complement each other. We code whether a speech contains an informational or critical question (coded as 1) or not (coded as 0). Questioning is an additional measure of engagement.⁶⁰

'Story-telling'. According to Polletta and Lee, 'story-telling' is the most important component of alternative forms of communication. In order to capture 'story-telling' empirically, we measure whether participants use personal narratives or experiences.⁶¹

For the construction of our measure, we leave one crucial component of deliberative quality aside, namely participation equality,⁶² since this does not really capture the formal quality of

⁵⁸ Burkhalter, Gastil and Kelshaw (2002), p. 403.

⁵⁹ See Cohen 1989.

⁶⁰ Stromer-Galley (2007), p. 12.

⁶¹ Polletta and Lee 2006.

⁶² See Thompson 2008.

arguments.⁶³ Moreover, participation equality requires a group-level rather than a speech-level analysis.⁶⁴ While we acknowledge the importance of participation equality as a standard of deliberative quality, space considerations require us to concentrate on aspects of formal argumentative quality and argumentative reciprocity.

Setting Thresholds for High and Low Deliberative Quality

Recent years have witnessed an increasing demand for setting one or more ‘threshold’ values for high and low deliberative quality.⁶⁵ In this article, we make a first attempt at setting thresholds for the various DQI indicators. Since our empirical analysis will only comprise a limited number of discussion groups, we set thresholds for high and low deliberative standards at the level of individuals. The threshold-level problem is intertwined with a level of analysis problem. To date, the quality of deliberation had only been checked at the level of individual speeches. But this is problematic: in order to achieve an overall maximum score, every speaker would not only have to justify their demands and arguments thoroughly in every single speech, they would also have to be simultaneously orientated towards the common good and be respectful at all times. Even staunch advocates of deliberation might agree that this is conceptually impossible, ignoring ‘economies of speech’ and the fact that in good conversations, arguments are not repeated all the time. Therefore, we have applied a *holistic approach* which analyses the overall deliberative performance of each speaker in an *entire discussion*.

To identify high and low quality standards for the DQI indicators, we draw from a classic conception of (overall) deliberative quality inspired by Habermas as well as Gutmann and Thompson.⁶⁶ We acknowledge that even under fairly ideal conditions (e.g., those in DPs), deliberative *standards* always remain ‘regulative ideals’ which can never be fully achieved in practice.⁶⁷ But if we understand deliberative quality as a continuum that includes realistic criteria (such as sophisticated justification or explicit respect) that individuals can sometimes achieve, then those criteria, which may be derived from the critical and emancipatory underpinnings of Habermas’s discourse theory,⁶⁸ provide *guides for action* that real people can both strive towards and achieve. In other words, although the ideal standards cannot be achieved, there are worthy, realistic, ‘good enough’, ‘do-able’ action guides that can be achieved – versions of the ideal that are close enough to the ideal to satisfy the ethical demands of the real world.⁶⁹ With regard to classic deliberation, there is broad agreement in the literature that this type of communication entails complex reasoning and is geared towards finding common understanding and common values.⁷⁰ Translated to the DQI indicators, *classic deliberation means that participants offer sophisticated rationales, refer to the common good, show explicit respect towards other participants’ arguments as well as empathy to other groups, and question what others have claimed*. Consequently, all DQI indicators are dichotomized in accordance with these cut levels, i.e., the high quality categories are given a value of 1, whereas the other categories are re-coded as 0.⁷¹ Moreover, the expectation in classic

⁶³ Estlund (2000); for participation equality in EuroPolis, see Gerber (2015).

⁶⁴ Stromer-Galley (2007), p. 19.

⁶⁵ Dryzek (2007), p. 244.

⁶⁶ See Gutmann and Thompson 1996; Steiner et al. 2004.

⁶⁷ See Mansbridge et al. (2010), n. 3.

⁶⁸ We thank an anonymous reviewer for highlighting this point.

⁶⁹ We thank Jane Mansbridge for helping us to think through this complex issue.

⁷⁰ See Bächtiger et al. 2010.

⁷¹ In Table A1 in the online appendix, we detail how this is done for each DQI component.

deliberation is that the different quality standards occur *simultaneously*, i.e., good deliberators should ideally comply with all quality standards (the ‘unitary deliberator model’). The various standards are also *equivalent* to each other, i.e., no priority or differential weight is given to specific indicators of deliberative quality. Finally, the inclusion of story-telling into a measure of deliberative quality can give rise to two scenarios: if story-telling is aligned with the other components of classic deliberative quality, then it is a sort of ‘rhetorical addition’ that deliberators employ in order to make abstract reasoning more accessible. By contrast, if story-telling is not aligned with the other components of classic deliberative quality, then it might represent a distinct form of expression that is used by less skilled deliberators or by specific social and cultural groups (as it was originally imagined to be used by feminist critics of deliberation).

As mentioned before, it would be overly demanding to expect that ordinary citizens constantly reach the various deliberative standards in discussion. Thus, we attenuate the standards: we do not expect that citizens live up to classic deliberative standards all the time or even on average, but only expect that citizens achieve the various quality standards *at least once* in the discussion. We acknowledge that there are many other ways to set thresholds for high and low deliberative quality. Yet given critics’ focus on a classic conception as well as the latter’s excellent empirical performance (see next section), we decided to limit our analysis to this specific understanding of deliberative capacities.⁷²

Aggregation of the Components: Item Response Analysis

In order to explore whether the pre-defined standards of classic and Habermasian-inspired deliberation represent a latent variable of deliberative quality, we use Bayesian item response theoretic model (IRT). IRT was originally developed in psychology and educational science to measure latent psychological constructs.⁷³ More specifically, IRT enables researchers to reconstruct individuals’ intelligence from their response to different items. Accordingly, the probability that i gives a correct answer to j can be modelled as follows:

$$\text{Logit}(\text{Prob.}(i \text{ gives the correct answer to } j)) = \gamma_j (\beta_i - \alpha_j),$$

with β_i being the intelligence level of respondent i , α_j the difficulty level of item j and γ_j the discrimination parameter.

This equation assumes that the probability of a correct answer is given by the extent to which the degree of intelligence exceeds the difficulty of the question. Hence, the larger the difference between the degree of intelligence β_i and the difficulty of the question α_j , the higher the probability that a correct answer is given by respondent i . γ_j represents the impact of the latent dimension on the response, thus it is called the discrimination parameter. If $\gamma_j = 0$, there is no relationship between the identified latent dimension and the response category. In other words, the higher the discrimination parameter, the more the item differentiates between subjects.

This logic can be translated to our research purpose in a straightforward way. We can interpret deliberative ability similar to intelligence in educational science, namely how well

⁷² Of course, judging participants by their maximal performance does not say anything about the overall quality of their utterances. However, as aggregate statistics in the Appendix show (Table A6 and Table A7), the EuroPolis speakers were considerably more respectful than disrespectful and they were also much more likely to refer to the common good than to display self-interest. This, however, does not account for justification rationality: inferior justification was more present than sophisticated justification.

⁷³ Rasch 1980. Note that our model corresponds to a 2-parameter IRT model, which is less restrictive than Rasch’s model.

citizens are able to achieve the various standards of deliberative quality (justification rationality, respect, etc.). If deliberative ability is a latent and uni-dimensional construct, the item response functions of the various deliberative standards should display similar slopes for the discrimination parameters. Put differently, an improvement of one's ability level increases the probabilities of reaching a certain standard in a similar way. Compared to factor analytic methods, which also relate deliberative standards to the underlying dimension via differentiated loadings, IRT has some advantages: first, while conventional factor analysis only models the covariance of the item responses as product of the latent characters of items and individuals, IRT models the response as a function of the difference of the latent characters (difficulty and ability), which fits better intuitively.⁷⁴ Secondly and related to the first point, IRT also makes the realistic assumption that some components of deliberative ability may be more difficult to achieve than others, which can be modelled with the difficulty parameter. Thirdly, factor analytic methods are not appropriate for our dummy coded data since normally distributed error terms are assumed.⁷⁵

DATA

Research Setting: EuroPolis

We analyse the deliberative abilities of citizens in the context of EuroPolis, a pan-European deliberative poll which took place in Brussels in May 2009 and gathered a random sample of 348 people to discuss the topics of migration and climate change.⁷⁶ During the three-day event in Brussels, participants were randomly assigned to twenty-five small groups. The groups were created with random variations of the languages spoken. Each group included participants from two to five different nationalities. The discussions were simultaneously translated in all languages spoken in the respective small groups. The small group discussions were led by trained facilitators.

Since analysing deliberative processes empirically is a highly demanding and time-consuming affair, we refrained from analysing all twenty-five small groups. Rather, we took a purposive sample of thirteen groups and limited our analysis to the migration topic. We decided to focus on discussions between citizens of new EU member states (post-2004) and citizens of older states of the EU. Moreover, we also wanted to focus on discussions between citizens from Western European and Southern European countries. Both distinctions will enable us to test for 'cultural' distortions, by simultaneously holding the variation within the group of Western European participants as small as possible. The participants from Western Europe mainly originate from one of the founding states of the European Union. We excluded groups including participants from the United Kingdom and from Nordic countries, since this would introduce an additional dimension of variance with regard to experience with EU affairs and experiences with migration. With regard to socioeconomic and psychological variables, however, our subsample is largely comparable to the rest of the EuroPolis participants. In the Appendix, we provide details on the composition of each group in our subsample (Table A5) as well as a comparison between our subsample and all other EuroPolis participants (Table A8).

We coded every single speech act according to the updated DQI as presented above. All in all, we coded 944 speeches within thirteen groups. An inter-coder reliability test by three

⁷⁴ Jackman 2001.

⁷⁵ One important assumption of item response models is the local independence of items. A correlation analysis of residuals shows that our data do not violate this assumption.

⁷⁶ Isernia and Fishkin 2014.

independent coders showed respectable levels of agreement (see Appendix, Table A3 and A4).⁷⁷ After the coding, we extracted for each participant the best performance on each of the six DQI indicators and then dichotomized this information according to whether the participant reached the quality standard (see above and Appendix, Table A1). Given the transnational nature of the EuroPolis project, we were not able to code all speeches in their original language but had to rely on the translations instead.⁷⁸ In order to make sure that this factor does not confound our analyses, we introduced a control for translated speeches in our model. The variable, however, did not yield a statistically significant effect and for our variables of interest, results did not change (see Appendix, Table A9).

Operationalization and Analysis

At the core of our analysis is a latent construct of classic and Habermasian-inspired deliberative quality, obtained via IRT analysis. Details of the IRT analysis are presented in the results section. The various aspects of the ‘deliberative incompetence and distortion thesis’ will be evaluated as follows. Regarding the ‘unitary deliberator model’, we check whether the different components of deliberative quality – justification rationality, respect, empathy, and inquisitiveness – form a compound and uni-dimensional phenomenon at the level of deliberating citizens. This is done on the basis of the IRT analysis. Regarding socio-economic variables, we focus on gender (1 = female; 0 = male), age (measured in years), education (measured as the age at end of the education process), and working class (measured as self-positioning; 1 = working class; 0 = other). Two categorical variables allow us to distinguish between participants coming from Central and Eastern, Southern or Western European countries. Regarding psychological variables, we focus on political interest (measured on an eleven-point scale from 0 (‘not at all’) to 10 (‘passionately’)), salience (seriousness of the immigration problem, measured on a scale from 0 (‘no problem at all’) to 10 (‘the most serious problem we face’)), and prior knowledge (number of correct answers to three knowledge questions on immigration: definition of a ‘Blue card worker’, the current form of the EU immigration policy and some figures on the EU’s immigrant population). In the next section, we focus on deliberative influence and link justification rationality at the group level to the participants’ opinion change, and provide details of this analysis.

RESULTS

General Deliberative Abilities and the Unitary Deliberator Model

First, we explore whether an ideal citizen deliberator exists in the real world. Using IRT, we check whether the different DQI components – justification rationality, respect, etc. – form a latent dimension of deliberative quality. It is well known that IRT suffers from an identification problem.⁷⁹ To identify the model, we apply a Bayesian approach with prior information. More specifically, we set a truncated normal distribution for the discrimination parameter of an arbitrary selected item in order to identify the direction of the underlying dimension.

⁷⁷ Notice, however, that for two indicators, level and content of justification, a first reliability test yielded only satisfactory results. Consequently, we clarified the disagreements, re-coded the data, and then performed a second inter-coder reliability test for level and content of justification. This test then yielded a very good reliability for the two indicators.

⁷⁸ All participants spoke in their mother tongue. We were able to code the original transcripts when participants’ mother tongue was English, French or German. For the rest, we had to rely on translations instead.

⁷⁹ Jackman 2001.

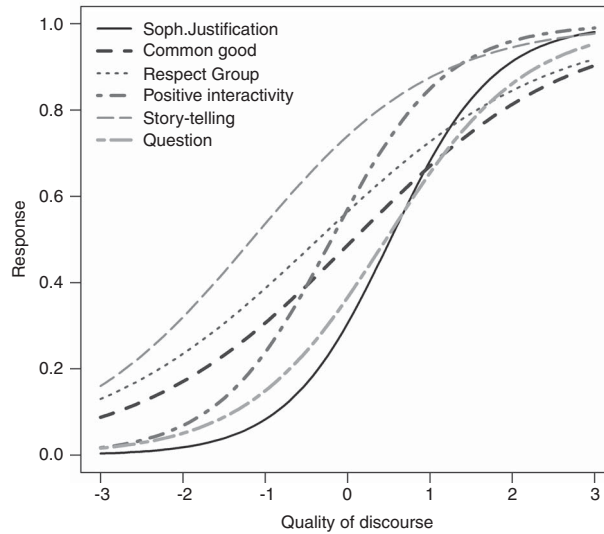


Fig. 1. Response functions for each item

Furthermore, we set a standard normal distribution for the ability parameter to set a certain scale for the underlying dimension. We ran three Markov chains with 7,000 iterations from different initial values, respectively. We discarded the first 2,000 iterations of each chain as burn-in and assessed convergence by visual inspection as well as by using the improved Brooks–Gelman–Rubin convergence diagnostics. We did not detect any sign of non-convergence.⁸⁰

Figure 1 presents the item response functions relating the latent dimension to the response probability of each indicator of deliberative quality. We see that all item response functions have similar positive slopes for the discrimination parameters, suggesting that all indicators are consistently related to a latent construct. Put differently, participants achieving the more difficult items (sophisticated justification) also have a higher probability of achieving the easier ones (e.g., respect towards other participants' arguments). This means that against scepticism from personality psychology, an ideal deliberator scoring high on justification rationality, common good orientation, respect, empathy and inquisitiveness exists in reality. The relatively steep discrimination parameters also indicate that participants with greater deliberative ability differ considerably from participants with less deliberative ability. A closer look at the difficulty parameters indicates that the demanding standards of classic deliberation were far from being rare events in the EuroPolis discussions. Zero on the horizontal axis corresponds to the mean of participants reaching this standard. We see that story-telling was the easiest item, with a 70 per cent chance that average-level EuroPolis participants reached this standard. For common good orientation, explicit respect towards other participants' arguments and explicit respect towards third-country migrants, there is about a 50 per cent chance that average-level participants reached these standards. Sophisticated justification was the most difficult item, with a 37 per cent chance that average-level participants reached this standard. An intriguing result is that story-telling – even though it represents the easiest standard – has a strong relationship with the latent dimension. This means that story-telling is a partial complement to justification

⁸⁰ The *Bayesian Item Response Analysis* was performed in WinBUGS, a free software package for Bayesian estimation. Information on the convergence diagnostics is available on request.

rationality, i.e., people who make sophisticated justifications also use story-telling. This also provides a hint that the classic distinction between rational discourse and alternative forms of communication may be misleading, since high-skilled deliberators also use personal experiences to back up their positions and arguments. This finding supports Ryfe's conclusions that 'successful deliberation seems to require a form of talk that combines the act of making sense (cognition) with the act of making meaning (culture). Storytelling is one such form of talk'.⁸¹

In sum, the results are quite striking: the standards of classic deliberation are far from being utopian standards that only very few citizen deliberators can achieve, as social psychologists have argued. The claim that ordinary citizens only make unelaborated statements and do not engage with each other is refuted by our data.⁸² Surely, the amount of 'deliberative all-rounders' is not large, but not extremely tiny either: focusing on the raw data, the number of participants reaching all six standards is 10 per cent; and if we consider those participants providing sophisticated justifications and simultaneously engaging in respectful listening, the amount goes up to almost 28 per cent.

Figure 2 documents variations of high and low deliberative quality across individuals and discussion groups. Notice that scores on the latent variable cannot be interpreted in an absolute way; but the more we move to the right-hand side of the continuum, the higher the deliberative quality (and vice versa). Individual points with 90 per cent credible intervals correspond to individual participants nested in thirteen discussion groups (we use the EuroPolis group classifications to denote the various groups).⁸³ Figure 2 clearly displays that all discussion groups involve participants with high and low deliberative ability. By the same token, there are no outstanding differences in deliberative ability among the thirteen discussion groups.

Exploring Socio-Economic, Cultural and Cognitive Distortions in Deliberative Abilities

In a second step, we explore whether particular socio-economic, cultural or psychological variables predict individual deliberative ability. The dependent variable is the latent deliberative ability estimated by IRT. We ran multilevel analyses to take into account that participants in EuroPolis are nested in discussion groups.⁸⁴ We include fixed effects for the different predictors and variance components at the levels of groups and individuals (i.e., we estimate a random intercept model).

⁸¹ Ryfe 2005, 63.

⁸² Rosenberg 2014.

⁸³ Credible intervals are the Bayesian analog to confidence intervals in conventional frequentist statistics. In accordance with the Bayesian approach, we should not interpret the overlapping credible intervals as the absence of a statistically significant difference, but the magnitude of overlap as the probability of indifference of ideal points. In other words, the less overlap the higher the probability that participants differ in their ability to deliberate. Participants are further sorted based on their individual deliberative qualities.

⁸⁴ With only thirteen discussion groups we face a small *n*-problem at level-2 units. Yet Gelman and Hill (2007, p. 247) argue that multilevel analysis is suitable with as few as five level-2 units. Most importantly, multilevel analysis does not perform worse than standard regression analysis, whereby the latter represents the limiting case of the former (Gelman and Hill 2007, p. 275). For parameter estimation, we utilized restricted maximum likelihood (REML) which considerably reduces bias in the context of a small number of groups (see for more detail Elff and Shikano 2014). We have re-checked our results by running standard regression analyses (using clusters at the group level). Results are almost identical compared to those reported in the multilevel estimations, albeit the latter yield more conservative results (analyses available upon request).

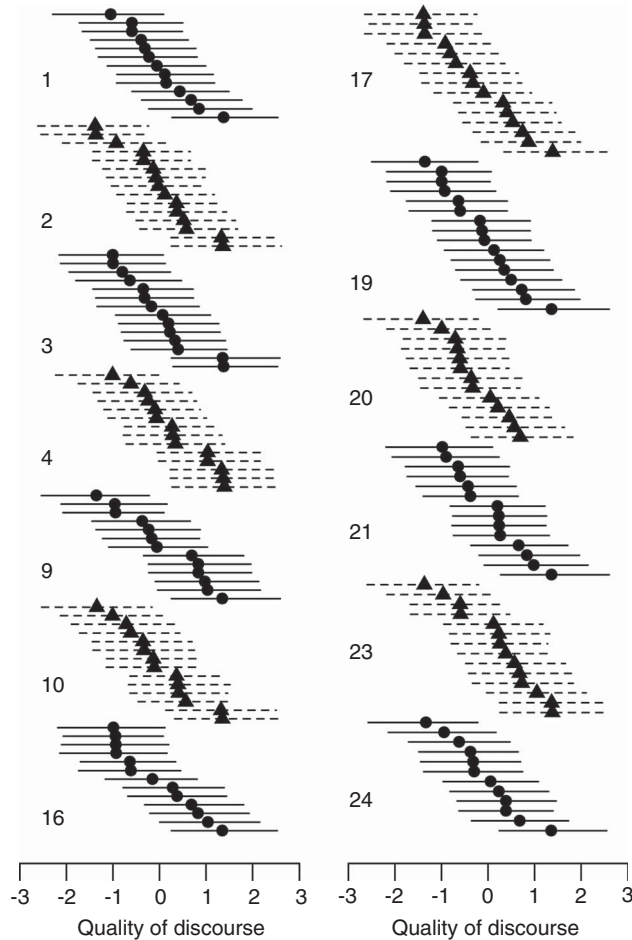


Fig. 2. Individual latent deliberative quality (sorted by group)

As Model 1 (Table 1) displays, working-class participants score lower on deliberative ability in terms of speaking skills than participants from higher classes ($p < 0.01$). This confirms sociologists arguing that working-class people possess a *repertoire* of speaking and arguing which may not so easily align with classic and Habermasian-inspired forms of deliberation.⁸⁵ We have also probed for various interaction effects among socio-economic and cultural variables, and found two (Model 2): working-class participants from both Central and Eastern Europe and Southern Europe (Spanish and Portuguese) do not reach the same levels of deliberative quality as other participants. Both interactive effects are substantively large and statistically significant ($p < 0.01$ and $p < 0.05$). This means that less privileged people in the European ‘polis’ – lower-class participants from the European periphery – were also the least skilled deliberators. Indeed, when focusing on the raw data, we see that working-class participants from Central and Eastern Europe, for instance, did not provide a single sophisticated rationale in the entire discussion. From a distortion perspective, this is a

⁸⁵ Note that in this article, we do not draw conclusions on the inner process of reasoning of working-class (or any other) participants.

TABLE 1 *Antecedents of Deliberative Quality*

	Model 1	Model 2
Gender	-0.106 (0.124)	-0.164 (0.123)
Education	0.009 (0.014)	0.006 (0.013)
Working class	-0.390** (0.148)	0.025 (0.199)
Age	-0.007 [†] (0.004)	-0.007 [†] (0.004)
Southern Europeans	-0.205 (0.160)	-0.048 (0.182)
Central Eastern Europeans	-0.230 (0.163)	-0.010 (0.176)
Saliency	-0.047 [†] (0.025)	-0.047 [†] (0.025)
Knowledge	0.091 (0.067)	0.117 [†] (0.066)
Interest in politics	0.054 [†] (0.030)	0.052 [†] (0.029)
SE × working class		-0.685* (0.341)
CEE × working class		-1.046** (0.358)
Constant	0.327 (0.475)	0.359 (0.464)
Std. Group Level	0.000*** (0.000)	0.000*** (0.000)
Std. Individual Level	0.745*** (0.043)	0.727*** (0.042)
N Individuals	163	163
N Groups	13	13

Note: Multilevel linear models (REML) with standard errors in parentheses.

[†] $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

worrisome finding: it means that already disadvantaged people have trouble adapting to deliberative modes of interaction, turning pan-European citizen deliberation into a ‘fragmented’ exercise privileging some and excluding others.

Notice, however, that Central and Eastern and Southern Europeans (Spanish and Portuguese) *per se* do not perform worse than Western Europeans. Region turns into a significant predictor only when it is interacted with class. Thus, at a general level, popular claims that culture is a powerful predictor of deliberative quality are not corroborated. Nonetheless, one may still wonder whether the two interaction effects are the product of different ‘speech cultures’ (with working-class people from Central and Eastern and Southern Europe having a different way of speaking and arguing than other participants), or whether they represent a ‘newcomer’ effect, especially mirroring the inexperience of Central and Eastern Europeans with pan-European affairs.⁸⁶ Of course, no definite answer can be given here. But the fact that the substantive effect is clearly stronger for Central and Eastern than for Southern Europeans which have a longer

⁸⁶ See Naurin 2010.

history of EU membership than Eastern Europeans provides a first hint that the effect may be more due to newcomer status than due to different ‘speech cultures’. Next, we find that higher salience levels lead to lower deliberative quality. This effect, however, is only marginally significant ($p < 0.10$). Nonetheless, this is an intriguing result, supporting classic deliberation’s claim that it is dispassionate attitudes – and not passionate ones – which are conducive to higher deliberative quality. Furthermore, and in accordance with our expectations, we find a slight tendency for higher political interest (a proxy for general motivation to engage in discussion) to be positively associated with higher deliberative quality ($p < 0.10$). As expected, there is also a slight positive association of prior knowledge with higher deliberative quality ($p < 0.10$ in the second model). Contrary to our expectations, age is slightly negatively associated with higher deliberative quality ($p < 0.10$). No statistically reliable effects were found for gender and education. Given their (enduring) prominence in the literature, the absence of gender effects is an important result.⁸⁷ It defies feminists’ claims and empirical findings in psychological experiments that women are almost always disadvantaged in deliberative processes. Of course, the absence of relevant differences does not mean that gendered patterns of communicating were absent, since masculine norms may have impinged on speaking styles with women adapting to these masculine norms. In the case of education, the absence of any effect seems surprising at first glance. However, in the EuroPolis case, we suspect that this result may be due to the particular way the EuroPolis questionnaire measures education – namely as ‘years of education’. This operationalization does not properly disentangle participants having higher education and a university degree from those who have not; yet this distinction may be the driving factor behind high and low deliberative quality.⁸⁸

In sum, citizen deliberation in supportive institutional environments such as deliberative polls works better than its most fervent critics have postulated. Not only is deliberation a more widespread ability than commonly assumed, the various components go together. This indicates that ideal ‘deliberative citizens’ simultaneously scoring high on justification rationality and respectful listening exist. By the same token, important socio-economic variables such as gender are not associated with deliberative quality. Nonetheless, citizen deliberation in deliberative polls is not immune from distortions, documented by the important class–region effect in EuroPolis. Focusing on perception-based measures, there is a negative correlation of perceived ‘self-silencing’ with our latent variable of deliberative ability (Pearson’s $r = -0.28$; $p < 0.00$), indicating that more highly skilled deliberators felt that they had better opportunities to make themselves heard during the deliberative proceedings.⁸⁹ The correlation is not particularly strong, but it still indicates that some voices of less skilled deliberative participants were not heard during the EuroPolis proceedings.

Deliberative Influence

In the theoretical section, we have argued that if opinion change is not associated with an argument-based pathway, then opinion change is normatively questionable. We have seen in

⁸⁷ See Karpowitz, Mendelberg and Shaker 2012. However, analysing online deliberative polls, Siu (2009) did not detect a gender effect for justification rationality either.

⁸⁸ Supporting this interpretation, ‘working class’ has no statistically significant association with this measure of education (Pearson’s $r = -0.18$, $p = 0.27$). If the education measure had properly separated people with higher education/university degree from others, then we should see a much stronger association of education with class. We also probed for an education indicator with dummies but this made no difference to the results.

⁸⁹ The exact wording is: ‘On a 0 to 10 scale, where “0” means “completely disagree”, “10” means “completely agree”, and “5” is “exactly in the middle”, how strongly would you agree or disagree with each of the following statements? “I kept some of my thoughts to myself for fear of the reactions of others.”’

the previous section that there were a considerable number of citizens putting forward well-justified arguments, which could serve as a *group resource* driving opinion change. With regard to argument-based opinion change, Sanders has performed such an analysis for EuroPolis by relying on participants' self-perceptions of discussion quality. Yet, he could not identify any significant effects.⁹⁰ In the following, we replicate Sanders study by using the data that we gathered via external DQI coding using justification rationality as a key indicator of deliberative quality.⁹¹ Even though we do not see justification rationality as the only normatively desirable pathway to deliberative opinion change, we focus on this specific indicator, for four reasons. First, at the level of group analysis, the various deliberative components do not constitute a conceptual whole (as was the case at the level of individuals). Therefore, we refrained from collapsing the different components into a single indicator of deliberative quality. Secondly, and related to the first reason, individual criteria can have different mechanisms in influencing the opinion so that the mix of multiple items can cancel out their impacts. Thirdly, justification rationality is not only conceptually a fairly straightforward criterion of deliberative quality, our empirical analysis shows that it also formed the most discriminating item at the individual level. Fourthly, an ANOVA test also reveals a significant variation of justification rationality among discussion groups ($p < 0.001$; by contrast, there were no significant differences for respect towards other participants' arguments).⁹²

Following Sanders, we expect attitude change to be greater in groups where the deliberative quality was higher, and *vice versa*.⁹³ Following Sanders, we conceptualize deliberative quality (justification rationality) as a group resource and operationalize it as the *mean performance of all the speeches in each group*. With regard to opinion change – our dependent variable – we focus on immigration attitudes. Like Sanders, we constructed a pro-immigration index based on twelve questions capturing respondents' attitudes towards third-country immigrants. Apart from the DQI-based operationalization of deliberative quality, we employ the same predictors as Sanders: gender, age, education, Catholic, Protestant, working-class, religiosity, left–right ideology (including a squared term), knowledge change, social conformity pressure, the intention to vote for left or right party groupings at the European Parliament elections and four questions asking whether experts, politicians, other participants or the briefing material helped to clarify thinking (see Table A10 in the Appendix for details on operationalization).⁹⁴

We estimated two models of opinion change where we compare post-deliberative opinions (wave 3; immediately after the DP event) to pre-deliberative opinions (wave 2; at the very beginning of the DP event).⁹⁵ In the first model, we model opinion change by taking immigration attitudes at wave 3 as a dependent variable, controlling for immigration attitudes

⁹⁰ Sanders 2012. The self-perception measure included several criteria – among others questions on whether the small group members participated equally, whether opposing arguments were considered, whether there were ample opportunities to express one's view, whether participants were respectful towards each other, and whether people offered justifications to back their views.

⁹¹ See Siu (2009) for a similar approach.

⁹² Storytelling ($p < 0.05$), common-good orientation and positive respect towards migrants (both $p < 0.001$) also showed significant ANOVA tests.

⁹³ Sanders (2012), p. 622.

⁹⁴ There are small differences in results compared to Sanders which stem from our focus on waves 3 and 2, particularly for knowledge change (see Appendix, Table A10). We also performed a series of robustness checks – including variables from our models on deliberative abilities (regional affiliation, interest, salience) but the results do not change.

⁹⁵ Because Sanders's main interest was comparing opinion shifts between the test and control group, he focused on opinions at wave 1 and wave 4 where test and control group could be compared.

TABLE 2 *Determinants of Opinion Change on Immigration Position (Replicating Sanders)*

	Immigration position (wave 3)	Absolute opinion change (immigration)
<i>Individual Level Determinants</i>		
Controls included	Yes	Yes
Immigration position (wave 2)	0.962*** (0.119)	
Change in knowledge (waves 2–3)	0.076 [†] (0.039)	–0.012 (0.018)
Clarify thinking – other participants	0.029 (0.020)	–0.001 (0.009)
Clarify thinking – experts	0.017 (0.024)	0.008 (0.011)
Clarify thinking – politicians	0.010 (0.021)	0.004 (0.010)
Clarify thinking – briefing material	–0.041 [†] (0.022)	0.000 (0.010)
Social conformity pressure (above)	0.003 (0.126)	–0.018 (0.024)
Social conformity pressure (below)	0.089 (0.130)	0.035 (0.023)
Constant	–0.049 (0.788)	0.219 (0.173)
<i>Group Level Determinants</i>		
Deliberative quality – level of justification	0.008 (0.178)	0.130* (0.066)
Std. Group Level	0.095 (0.064)	0.000 (0.000)
Std. Individual Level	0.455 (0.027)	0.211 (0.012)
N Individuals	174	174
N Groups	13	13

Note: Multilevel linear models (REML) with standard errors in parentheses. Controls included (full model reported in Appendix, Table A11) for gender, age, education, Catholic, Protestant, working class, religiosity, left–right ideology, left–right squared, intention to vote for left (PES, far left or green) or right party (EPP, far right or Libertas) at the 2009 elections to the EU Parliament. For operationalization, see Table A10 in the Appendix.

[†] $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

at wave 2.⁹⁶ Besides the direction, we also focus on the *magnitude* of opinion change between wave 2 and 3. Since the absolute amount of opinion change is right-skewed, we operationalized the magnitude of change as the logarithm of absolute opinion change. For both models, we calculated linear multilevel regressions with individuals nested in groups.⁹⁷

The first model in Table 2 shows that justification rationality does not affect directional shifts of immigration attitudes. This may not be so surprising if one considers that justification

⁹⁶ See Sanders 2012.

⁹⁷ Sanders (2012) estimated linear regressions with clustered standard errors for group affiliation. In contrast, we include the effect of individual groups as random intercept and the group-level predictors (as average-level of deliberative quality). Therefore, we estimated multi-level models.

rationality is only a formal quality criterion which does not dictate any specific direction of opinion change.⁹⁸ By contrast, justification rationality has a clear effect on the magnitude of opinion change: the higher the group level of justification, the more participants have reconsidered their original positions on immigration (Model 2b).⁹⁹ This important result is in full accordance with deliberative theory. Moreover, this normatively desirable effect is bolstered by the fact that there are no signs of social conformity pressures: participants did not adapt their post-deliberative opinions to the average of the pre-deliberative opinions in the group.

Finally, let us consider potential distortions by high-skilled deliberators on opinion change. While our previous analysis shows that high-skilled deliberators are not only good at providing sophisticated justifications but also listen respectfully, they do not seem to impose their original opinions on other participants either. Highly skilled deliberators were not stuck in their positions: compared to the 90 per cent of participants with less skillful deliberative abilities, they showed an almost identical amount of absolute opinion change on our immigration scale after deliberations ended ($t = 0.19$). Furthermore, as a supplementary analysis shows (see Appendix, Table A12), the *pre-deliberative* opinions of highly skilled deliberators did not affect other participants' *post-deliberative* opinions.¹⁰⁰ While tentative, these results provide a further indication that while opinion change in EuroPolis is affected by argument, the effect does not seem to be produced by 'authority',¹⁰¹ in the form of imposed argumentative influence by deliberatively-skilled participants.

CONCLUSION

This is one of the few studies exploring deliberative abilities of ordinary citizens on the basis of a philosophically grounded measure, an updated version of the Discourse Quality Index (DQI). The object of our study was EuroPolis, a transnational deliberative poll. The EuroPolis event yields a mixed but ultimately fairly optimistic picture of citizen deliberation, defying many allegations made in the context of the 'deliberative incompetence and distortion thesis'. First, citizens' deliberative abilities are more widespread than assumed by sceptics: demanding standards of classic deliberation (such as sophisticated justifications and respectful listening) are far from being utopian standards that only a tiny minority of citizen deliberators can achieve. Moreover, the ideal deliberator scoring high on all deliberative dimensions as envisaged by classic deliberative theory exists. An item response analysis shows that classic deliberative ability forms a latent dimension with participants scoring high on all deliberative standards, ranging from justification rationality to common good orientation, respect, empathy and inquisitiveness. Secondly, while high deliberative ability is only weakly correlated with social and cognitive characteristics, we nonetheless found that less privileged people in the European 'polis' – lower-class participants, particularly from the European periphery – were also the least skilled deliberators. Compared to other participants, working-class participants from Central and Eastern as well as Southern Europe were much less likely to reach the various standards of high-quality deliberation, raising some concerns about the democratic dimensions of

⁹⁸ For a slightly different approach, i.e. combining justification rationality with arguments favouring and opposing an issue, see Siu (2009) and Gerber et al. (2014).

⁹⁹ For similar results see Schneiderhahn and Khan (2008) and Siu (2009).

¹⁰⁰ This resonates with findings from Fishkin, Luskin and Siu (2014): their findings on social influence in EuroPolis do not indicate any domination by the more advantaged (men and upper class). See also Siu (2009) and Fishkin et al. (2010) for similar results on US and Chinese deliberative polls.

¹⁰¹ See Karpowitz, Mendelberg and Shaker 2012.

deliberation among citizens with heterogeneous backgrounds. Thirdly, there is evidence that the higher the justification rationality within a discussion group, the more participants changed their opinions. This is an important finding (especially in the light of Sanders's daunting results for opinion change in EuroPolis), showing that argument and not other (non-deliberative) dynamics drive opinion formation.¹⁰² In addition, we find no evidence that highly skilled deliberators were not receptive to other participants' claims, closed-minded, or imposed their pre-deliberative opinions on other participants. Of course, our optimistic results need to be replicated; but the fact that these results were obtained under 'difficult conditions', namely in the context of transnational citizen deliberation, provides a strong indication that deliberative polls seem to work better than critics and sceptics of deliberation have claimed.

Of course, our study is not without limitations. First, one might object that we have only analysed thirteen out of twenty-five discussion groups in EuroPolis. However, since there are no signs of systematic variation across discussion groups, we wonder how much empirical leverage will be gained by analysing more groups. Secondly, some of the measures used in the analysis are not optimal and stronger associations might have been found if we had other measures, as for instance psychological factors. There is new research linking personality traits to electoral behaviour, political participation and political attitudes.¹⁰³ There is an urgent need to include such variables in future questionnaires of citizen deliberation as well. Thirdly, our analysis only considers the *formal* aspects of highly deliberative quality talk while neglecting other aspects such as the *substantive* content of arguments or argumentative balance.¹⁰⁴ We believe that it is important to distinguish formal from substantive and issue-specific contents. This would answer the question why different studies obtain such widely differing results on inequalities in deliberation. Fourthly, our main focus lies on highly deliberative quality *talk*. Analysing the speaking behaviour of EuroPolis participants deprives us from knowing whether the participants also undertook 'deliberation within'.¹⁰⁵ However, our findings on opinion formation offer a first indication that considered weighing of arguments might have taken place. Fifthly, deliberative polls aim at simulating ideal conditions for successful deliberation and deprive participants from reaching a decision or a common statement in the end in order to 'insulate people from social pressure'.¹⁰⁶ This prevents us from translating our findings to real-world public deliberation. More and systematic research is required to analyse the deployment of citizens' deliberative abilities, as well as the deliberative quality as a whole, in various contexts. Finally, our study is only concerned with the 'internal quality' of deliberative mini-publics, not with its external one.¹⁰⁷ But even if the internal and external quality may not always overlap in practice, a trust-based 'uptake' of mini-public recommendations¹⁰⁸ seems highly questionable when the internal proceedings of a deliberative mini-public do not work according to (or, against) deliberative principles.

Despite these limitations, our study is the first one to perform a comprehensive, systematic and in-depth analysis of the deliberative abilities of ordinary citizens and deliberative influence in a deliberative poll. At the same time, it provides researchers with a tool set that can be applied to an in-depth evaluation of the booming numbers of citizen deliberative events world-wide.

¹⁰² E.g., Barabas 2004; Chambers (1996), p. 99; Neblo 2007.

¹⁰³ See Mondak et al. 2010.

¹⁰⁴ See Fishkin (2009), p. 34.

¹⁰⁵ Goodin 2000.

¹⁰⁶ Ackermann and Fishkin (2002), p. 134.

¹⁰⁷ See Curato and Boeker 2015.

¹⁰⁸ MacKenzie and Warren 2012.

REFERENCES

- Ackerman, Bruce, and James S. Fishkin. 2002. Deliberation Day. *Journal of Political Philosophy* 10:129–152.
- Andersen, Vibeke N., and Kasper M. Hansen. 2007. How Deliberation Makes Better Citizens: The Danish Deliberative Poll on the Euro. *European Journal of Political Research* 46:531–56.
- Bächtiger, André, Simon Niemeyer, Michael Neblo, Marco R. Steenbergen, and Jürg Steiner. 2010. Disentangling Diversity in Deliberative Democracy: Competing Theories, Their Blind Spots and Complementarities. *Journal of Political Philosophy* 18:32–63.
- Bernstein, Basil. 1971. *Class, Codes and Control: Theoretical Studies towards a Sociology of Language*. London: Routledge & Kegan.
- Burkhalter, Stephanie, John Gastil, and Todd Kelshaw. 2002. A Conceptual Definition and Theoretical Model of Public Deliberation in Small Face-to-Face Groups. *Communication Theory* 12: 398–422.
- Caluwaerts, Didier. 2012. *Confrontation and Communication: Experiments on Deliberative Democracy in Linguistically Divided Belgium*. Brussels: Vrije Universiteit Brussel.
- Cappella, Joseph N., Vincent Price, and Lilach Nir. 2002. Argument Repertoire as a Reliable and Valid Measure of Opinion Quality: Electronic Dialogue during Campaign 2000. *Political Communication* 19:73–93.
- Chambers, Simone. 1996. *Reasonable Democracy: Jürgen Habermas and the Politics of Discourse*. Ithaca, N.Y.: Cornell University Press.
- . 2004. Behind Closed Doors: Publicity, Secrecy, and the Quality of Deliberation. *Journal of Political Philosophy* 12:389–410.
- Clinton, Joshua, Simon Jackman, and David Rivers. 2004. The Statistical Analysis of Roll Call Data. *American Political Science Review* 90:355–70.
- Cohen, Joshua. 1989. ‘Deliberation and Democratic Legitimacy’. Pp. 17–34 in *The Good Polity*, edited by Alan Hamlin and Philip Pettit. Oxford: Blackwell.
- Curato, Nicole, and Marit Böker. 2015. Linking Mini-Publics to the Deliberative System: A Research Agenda. *Policy Sciences*, Online First.
- Doerr, Nicole. 2009. Activists beyond Language Borders? Multilingual Deliberative Democracy Experiments at the European Social Forums. 5th ECPR General Conference, Potsdam.
- Dryzek, John S. 2007. Theory, Evidence, and the Task of Deliberation. Pp. 237–50 in *Can the People Govern? Deliberation, Participation and Democracy*, edited by Shawn Rosenberg. New York: Palgrave Macmillan.
- . 2010. *Foundations and Frontiers of Deliberative Governance*. Oxford: Oxford University Press.
- Duchesne, Sophie, and Florence Haegel. 2007. Avoiding or Accepting Conflict in Public Talk. *British Journal of Political Science* 37:1–22.
- Dutwin, David. 2003. The Character of Deliberation: Equality, Argument, and the Formation of Public Opinion. *International Journal of Public Opinion Research* 15:239–64.
- Elff, Martin, and Susumu Shikano. 2014. Bias in ML Estimation of Multilevel Models: Check the Algebra before You Gamble in Monte Carlo. Paper presented at the 4th Annual General Conference of the European Political Science Association, Edinburgh, 19 = 21 June 2014.
- Estlund, David. 2000. Political Quality. *Social Philosophy and Policy* 17:127–60.
- Fishkin, James S. 1995. *The Voice of the People: Public Opinion and Democracy*. New Haven, Conn.: Yale University Press.
- . 2009. *When the People Speak: Deliberative Democracy and Public Consultation*. Oxford: Oxford University Press.
- Fishkin, James S., Baogang He, Robert C. Luskin, and Alice Siu. 2010. Deliberative Democracy in an Unlikely Place: Deliberative Polling in China. *British Journal of Political Science* 40: 435–448.
- Fishkin, James S., Robert C. Luskin, and Alice Siu. 2014. EuroPolis and the European Public Sphere: Empirical Explorations of a Counterfactual Ideal. *European Union Politics* 15:328–351.
- Fuchs, Dieter, and Hans-Dieter Klingemann. 2002. Eastward Enlargement of the European Union and the Identity of Europe. *West European Politics* 25:19–54.

- Fung, Archon. 2003. Survey Article: Recipes for Public Spheres: Eight Institutional Design Choices and Their Consequences. *Journal of Political Philosophy* 11:338–67.
- Gambetta, Diego. 1998. Claro! An Essay on Discursive Machismo. Pp. 19–43 in *Deliberative Democracy*, edited by John Elster. Cambridge: Cambridge University Press.
- Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.
- Gerber, Marlène. 2015. Equal Partners in Dialogue? Participation Equality in a Transnational Deliberative Poll (EuroPolis). *Political Studies* 63:110–30.
- Gerber, Marlène, André Bächtiger, Irena Fiket, Marco Steenbergen, and Jürg Steiner. 2014. Deliberative and Non-Deliberative Persuasion: Mechanisms of Opinion Formation in EuroPolis. *European Union Politics* 15:410–29.
- Goodin, Robert E. 2000. Democratic Deliberation Within. *Philosophy & Public Affairs* 29:81–109.
- . 2005. Sequencing Deliberative Moments. *Acta Politica* 40:182–96.
- Grönlund, Kimmo, Maija Setälä, and Kaisa Herne. 2010. Deliberation and Civic Virtue: Lessons from a Citizen Deliberation Experiment. *European Political Science Review* 2:95–117.
- Gutmann, Amy, and Dennis F. Thompson. 1996. *Democracy and Disagreement*. Cambridge, Mass.: Harvard University Press.
- Habermas, Jürgen. 2005. Concluding Comments on Empirical Approaches to Deliberative Politics. *Acta Politica* 40:384–92.
- Hansen, Kasper M. 2004. *Deliberative Democracy and Opinion Formation*. Odense: University Press of Southern Denmark.
- Hastie, Reid, Steven Penrod, and Nancy Pennington. 1983. *Inside the Jury*. Cambridge, Mass.: Harvard University Press.
- Himmelroos, Staffan, and Henrik Serup Christensen. 2014. Deliberation and Opinion Change: Evidence from a Deliberative Mini-Public in Finland. *Scandinavian Political Studies* 37: 1–41.
- Hooghe, Marc. 1999. The Rebuke of Thersites. *Deliberative Democracy under Conditions of Inequality*. *Acta Politica* 34:287–301.
- Isernia, Pierangelo, and James Fishkin. 2014. The EuroPolis Deliberative Poll. *European Union Politics* 15:311–27.
- Jackman, Simon. 2001. Multidimensional Analysis of Roll Call Data via Bayesian Simulation: Identification, Estimation, Inference, and Model Checking. *Political Analysis* 9:227–41.
- Jennstal, Julia. 2012. Traits and Talks: Lessons about Personality and Deliberation from the Negotiations between Nelson Mandela and F. W. De Klerk. doctoral dissertation, University of Uppsala, Sweden.
- Jennstal, Julia, and Simon Niemeyer. 2014. The Deliberative Citizen: Exploring Who is Willing to Deliberate, When and How through the Lens of Personality, Centre for Deliberative Democracy and Global Governance Working Paper Series, Working Paper No. 1, University of Canberra, Centre for Deliberative Democracy and Global Governance.
- Karp, Jeffrey A., and Susan A. Banducci. 2007. Party Mobilization and Political Participation in New and Old Democracies. *Party Politics* 13:217–34.
- Karpowitz, Christopher F., and Chad Raphael. 2014. *Deliberation, Democracy, and Civic Forums: Improving Equality and Publicity*. New York: Cambridge University Press.
- Karpowitz, Christopher F., Tali Mendelberg, and Lee Shaker. 2012. Gender Inequality in Deliberative Participation. *American Political Science Review* 106:533–47.
- Kim, Nuri, Alice Siu, and Gaurav Sood. 2010. Minority Report: Impact of Opinion Minorities on Deliberation. Paper presented at the Annual Meeting of the International Communication Association, Singapore, 21–26 June 2010.
- Landemore, Hélène, and Hugo Mercier. 2012. Talking It Out with Others vs. Deliberation Within and the Law of Group Polarization: Some Implications of the Argumentative Theory of Reasoning. *Análise Social* 205:910–34.
- Luskin, Robert C., Gaurav Sood, James S. Fishkin, and Kyu S. Hahn. 2015. Paper presented at the Congress of the French Political Science Association, Aix-en-Provence.

- Luskin, Robert C., Ian O'Flynn, James S. Fishkin, and David Russel. 2014. Deliberating across Deep Divides. *Political Studies* 62:116–35.
- Luskin, Robert C., James S. Fishkin, and Roger Jowell. 2002. Considered Opinions: Deliberative Polling in Britain. *British Journal of Political Science* 32:455–87.
- MacKenzie, Michael K., and Mark E. Warren. 2012. Two Trust-Based Uses of Minipublics in Democratic Systems. Pp. 95–124 in *Deliberative Systems: Deliberative Democracy at the Large Scale*, edited by John Parkinson and Jane Mansbridge. Cambridge: Cambridge University Press.
- Mansbridge, Jane. 2010. Deliberative Polling as the Gold Standard. *The Good Society* 19:55–62.
- Mansbridge, Jane, with James Bohman *et al.* 2010. The Place of Self-Interest and the Role of Power in Deliberative Democracy. *Journal of Political Philosophy* 18:64–100.
- Mansbridge, Jane *et al.* 2012. A Systemic Approach to Deliberative Democracy. Pp. 1–26 in *Deliberative Systems: Deliberative Democracy at the Large Scale*, edited by John Parkinson and Jane Mansbridge. Cambridge: Cambridge University Press.
- McCrae, Robert R., and Paul T. Costa. 1987. Validation of the Five-Factor Model of Personality across Instruments and Observers. *Journal of Personality and Social Psychology* 52:81–90.
- Mendelberg, Tali. 2002. The Deliberative Citizen: Theory and Evidence. Pp. 151–93 in *Political Decision Making, Deliberation and Participation: Research in Micropolitics*, Vol. 6, edited by Michael X. Delli Carpini, Leonie Huddy and Robert Y. Shapiro. Greenwich, Conn.: JAI Press.
- Min, Seong-Jae. 2009. Deliberation, East Meets West: Exploring the Cultural Dimension of Citizen Deliberation. *Acta Politica* 44:439–58.
- Mondak, Jeffery J., Matthew V. Hibbing, Damarys Canache, Mitchell A. Seligson, and Mary R. Anderson. 2010. Personality and Civic Engagement: An Integrative Framework for the Study of Trait Effects on Political Behavior. *American Political Science Review* 104:85–110.
- Mutz, Diana C. 2008. Is Deliberative Democracy a Falsifiable Theory? *Annual Review of Political Science* 11:521–38.
- Naurin, Daniel. 2010. Most Common When Least Important: Deliberation in the European Union Council of Ministers. *British Journal of Political Science* 40:31–50.
- Neblo, Michael A. 2007. Change for the Better? Linking the Mechanisms of Deliberative Opinion Change to Normative Theory. Available from polisci.osu.edu/faculty/mneblo/papers/ChangeC4.pdf, accessed 12 May 2015.
- Nie, Norman H., Jane Junn, and Kenneth Stehlik-Barry. 1996. *Education and Democratic Citizenship in America*. Chicago: University of Chicago Press.
- Norris, Pippa. 1996. Women Politicians: Transforming Westminster? *Parliamentary Affairs* 49:89–102.
- Pedrini, Seraina. 2014. Deliberative Capacity in the Political and Civic Sphere. *Swiss Political Science Review* 20:263–86.
- Petty, Richard E., John T. Cacioppo, and David Schumann. 1983. Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement. *Journal of Consumer Research* 10:135–46.
- Polletta, Francesca, and John Lee. 2006. Is Telling Stories Good for Democracy? Rhetoric in Public Deliberation after 9/11. *American Sociological Review* 71:699–723.
- Rasch, Georg. 1980. *Probabilistic Models for Intelligence and Attainment Tests*. Chicago: University of Chicago Press.
- Ratner, Steven R. 2008. Communicative Rationality in the Citizens' Assembly and Referendum Processes. Pp. 145–65 in *Designing Deliberative Democracy: The British Columbia Citizens' Assembly*, edited by Mark E. Warren and Hilary Pearse. Cambridge: Cambridge University Press.
- Rosenberg, Shawn W. 2005. The Empirical Study of Deliberative Democracy: Setting a Research Agenda. *Acta Politica* 40:212–24.
- . 2007. Types of Discourse and Democracy of Deliberation. Pp. 130–58 in *Can the People Govern? Deliberation, Participation and Democracy*, edited by Shawn W. Rosenberg. New York: Palgrave Macmillan.

- . 2014. Citizen Competence and the Psychology of Deliberation. Pp. 98–117 in *Deliberative Democracy: Issues and Cases*, edited by Stephen Elstub and Peter McLaverty. Edinburgh: Edinburgh University Press.
- Ryfe, David M. 2005. Does Deliberative Democracy Work? *Annual Review of Political Science* 8:49–71.
- Sanders, David. 2012. The Effects of Deliberative Polling in an EU-Wide Experiment: Five Mechanisms in Search of an Explanation. *British Journal of Political Science* 42:617–40.
- Sanders, Lynn M. 1997. Against Deliberation. *Political Theory* 25:347–76.
- Schneiderhan, Erik, and Shamus Khan. 2008. Reasons and Inclusion: The Foundation of Deliberation. *Sociological Theory* 26:1–24.
- Shikano, Susumu. 2008. The Dimensionality of German Federal States: Policy Preferences in the Bundesrat. *German Politics* 17:340–52.
- Siu, Alice. 2009. Look Who's Talking: Examining Social Influence, Opinion Change and Argument Quality in Deliberation. Doctoral dissertation, Stanford University, Department of Communication.
- Steiner, Jürg. 2012. *The Foundations of Deliberative Democracy: Empirical Research and Normative Implications*. Cambridge: Cambridge University Press.
- Steiner, Jürg, André Bächtiger, Markus Spörndli, and Marco R. Steenbergen. 2004. *Deliberative Politics in Action*. Cambridge: Cambridge University Press.
- Stromer-Galley, Jennifer. 2007. Measuring Deliberation Content: A Coding Scheme. *Journal of Public Deliberation* 3:1–35.
- Thompson, Dennis F. 2008. Deliberative Democratic Theory and Empirical Political Science. *Annual Review of Political Science* 11:497–520.
- University of Texas. 2003. *An Expertise in Democracy*. Available from <http://www.utexas.edu/features/archive/2003/polling.html>, accessed 27 February 2013.
- Verba, Sidney, Kay Lehman Schlozman, and Henry E. Brady. 1995. *Voice and Equality: Civic Voluntarism in American Politics*. Cambridge, Mass.: Harvard University Press.
- Young, Iris M. 2002. *Inclusion and Democracy* (paperback edition). Oxford: Oxford University Press.