

Article

Four Parameters for Measuring Democratic Deliberation: Theoretical and Methodological Challenges and How to Respond

Dannica Fleuß *, Karoline Helbig and Gary S. Schaal

Department of Political Science, Helmut-Schmidt-University, 22043 Hamburg, Germany; E-Mails: dannica.fleuss@hsu-hh.de (D.F.), karoline.helbig@hsu-hh.de (K.H.), gschaal@hsu-hh.de (G.S.S.)

* Corresponding author

Submitted: 30 September 2017 | Accepted: 8 November 2018 | Published: 19 March 2018

Abstract

Although measuring democratic deliberation is necessary for a valid measurement of the performance of democracies, it poses serious theoretical and methodological challenges. The most serious problem in the context of research on democratic performance is the need for a theoretical and methodological approach for “upscaling” the measurement of deliberation from the micro and meso level to the macro level. The systemic approach offers a useful framework for this purpose. Building on this framework, this article offers a modular approach consisting of four parameters for conceptualization, measurement, and aggregation which can be adjusted to make the measurement of democratic deliberation compatible with the various general measurement approaches adopted by different scholars.

Keywords

deliberation; democracy; democratic performance; measurement of democracy; systemic approach

Issue

This article is part of the issue “Why Choice Matters: Revisiting and Comparing Measures of Democracy”, edited by Heiko Giebler (WZB Berlin Social Science Center, Germany), Saskia P. Ruth (German Institute of Global and Area Studies, Germany), and Dag Tanneberg (University of Potsdam, Germany).

© 2018 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

For a long time, liberal democracies’ legitimacy mainly rested on voting and representation. In the course of the so-called “participatory revolution” (Kaase, 1984), the means of participation have increased dramatically in Western democracies. The term “democratic innovation” refers to the new, multi-faceted forms of participation which go beyond voting. Most of them are built around the idea and practice of deliberation in one way or another (Geißel & Newton, 2012). Today, the theory of deliberative democracy is considered to be the most important normative theory of democracy (Dryzek, 2015; Elstub, 2015). Accordingly, authors such as Dryzek (2010, pp. 21–42) argue that deliberative legitimacy has become the most important paradigm of legit-

imacy in contemporary political *theory* as well as democratic *practice*.

Due to the vastly increased theoretical importance and empirical impact of democratic theories of deliberation, measures of democracy need to include deliberation to achieve valid and empirically meaningful results. The Discourse Quality Index (DQI) already offers a sophisticated and widely acclaimed measuring instrument for democratic deliberation at the *micro/meso level*. However, an evaluation of the deliberative performance of democratic political systems at large requires measuring *deliberation at the macro level*. Niemeyer (2014) questioned whether scaling up deliberation was possible. A theoretically and methodologically grounded framework for this purpose is still required (Niemeyer, Curato, & Bächtiger, 2015, p. 4).

Addressing this gap in research, we develop an approach for measuring the deliberative performance¹ of political systems and systematically outline four “parameters for measuring macro deliberation” (PMMD). Thereby, we propose a modular approach that can be adopted by other scholars with different normative and theoretical presuppositions. The parameters are based on the assumption that the so-called “systemic approach”, as originally developed in a seminal book by Mansbridge et al. (2012), is the only framework for conceptualizing democratic deliberation² at the macro level to have been suggested so far which represents a suitable basis for measuring deliberative performance at this level. In this approach, deliberation is conceptualized as an “emergent property” (Niemeyer et al., 2015) which cannot be reduced to a mere aggregation of other qualities of the political system (see O’Conner & Wong, 2015). Rather than isolated deliberative fora, “the interdependence of sites within a larger system” as well as the interactions of deliberations in different institutions (and *loci* in general) represent the focal point of this understanding (Bohman, 2012, p. 73; Mansbridge et al., 2012, pp. 1f.).

In order to have a common point of departure, we take the original systemic approach as a starting point to develop a framework for “scaling up” the measurement of democratic deliberation from the micro to the macro level (see Niemeyer, 2014).³ This does not imply any commitment to Mansbridge et al.’s (2012) specific *normative* premises (and especially not to their concept of deliberation), but only to the basic framework of the systemic approach. Following an explanation of the need for a new approach to measure deliberation at the macro level in section 2, we outline the four “parameters” that have to be considered in the process of conceptualization and operationalization: *the theory of democracy, the concept of deliberation, the selection of loci, and the aggregation rule* (section 3). In the concluding section of this article, we identify some challenges future research will need to deal with regarding the measurement of the deliberative performance of democratic political systems.

2. Why “Parameters for Measuring Macro Deliberation”?

Elstub, Ercan and Mendonça (2016) identify four generations of the deliberative democracy school of thought. They started out with an explicitly normative theory on the rational, impartial justification of norms in the first generation (Cohen, 1989, 1997; Habermas, 1996),

continued to adapt the definition of deliberation to the increasing plurality and complexity of contemporary democracies in the second generation (Chambers, 2003, p. 322; Elstub et al., 2016, pp. 141f.), and began the empirical evaluation of deliberation under “laboratory conditions” in the third generation (Fishkin, 1995). In the following decades, “real world deliberation” became an object of scholars’ interest. With the DQI, the “gold standard” for the evaluation of institutions or the public sphere’s deliberative performance was developed (Mansbridge, 2010; Steiner, Bächtiger, Spörndli, & Steenbergen, 2004) and scholars such as Fung (2006, p. 66) discussed the “range of institutional possibilities for public participation”.

The *fourth generation* is characterized by the so-called “systemic turn”, which attempts to combine “the insights gained from three preceding generations, namely the strong normative premises, institutional feasibility, and empirical results” (Elstub et al., 2016, p. 143). The major innovation of the systemic approach is the acknowledgement of the “importance of looking at the system as a whole, as well as its different parts” rather than the previous focus on “isolated instances of deliberation” (Erman, 2016, p. 263). Thereby, “the deliberative system reconnects deliberative democratic theory to its initial macro ambitions: to enhance and understand democracy at the large scale” (Boswell & Corbett, 2017, p. 3). This implies that the systemic approach attempts to conceptualize democratic deliberation(s) as taking place all over a society or a political system and seeks to systematically account for the interactive relationships between various deliberative practices (Elstub et al., 2016, p. 140).

In spite of the considerations of *fourth generation* scholars, even two of the most sophisticated contemporary measures of democracy are unable to provide an appropriate measure of democratic deliberation at the macro level: while the *Democracy Barometer* (Merkel et al., 2016) does not take democratic deliberation into account at all, *Varieties of Democracy* (Coppedge et al., 2016) explicitly offers a “deliberative component index”. However, the latter demonstrates one of the major problems of addressing democratic deliberation in a democracy index: in their attempt to transfer criteria that were applicable at a micro/meso level to a larger scale, Coppedge et al. (2016) do not take account of the difference between the criteria for deliberation at the micro/meso level and at the macro level. Although they claimed to measure “deliberation at all levels” (Coppedge et al., 2016, p. 6), the items mostly address formalized deliberation by political elites. More-

¹ The concept used by John S. Dryzek (2009); Stevenson and Dryzek (2012) and many subsequent researchers (e.g. Felicetti, Niemeyer, & Curato, 2016; Niemeyer et al., 2015) is the concept of “deliberative capacity”. Although this concept undoubtedly “provides diagnostic criteria for assessing the system” (Felicetti et al., 2016, p. 429), we rely on the concept of “(deliberative) performance”, which is much more common in the context of measuring the performance of democratic political systems.

² We explicitly address the measurement of “democratic deliberation”, not “deliberative democracy”. While the latter denotes a political system that meets the demanding normative standards of deliberative theories, the latter refers to a form of communication, originally derived from these theories, put into practice in (liberal) democracies (Mansbridge, 2007).

³ More recent developments by Hendriks (2016a, 2016b), Ercan, Hendriks and Boswell (2017) and Erman (2016) will be integrated in section 3 of this article.

over, the aggregation rule is simply a factor index (that is, an additive index weighted by the factor loadings of the items) and does not reflect the complex interdependencies between those levels, especially not those between the different loci of deliberation. This shows that a valid measurement of democratic deliberation at the systems level “requires more than scaling up micropolitical concepts” (Niemeyer et al., 2015, p. 5). The different *kinds* of potential loci and deliberative practices, as well as the interactive relations, need to be taken into consideration and appropriately reproduced in the aggregation rule.

Thus, the systemic approach itself provides opportunities for research on democratic deliberation while addressing the “scaling-up” problem (Chambers, 2012; Elstub et al., 2016, p. 140; Erman, 2016, p. 263): the systemic approach conceptualizes the “deliberative quality” of a democracy as an “emergent property” of the system as a whole. This means that it is “irreducible” to the properties of parts of the system (the quality of deliberations in individual loci). Consequently, an aggregation rule that merely adds up the deliberative performances *within* those loci without taking account of the interactions of the “individual” deliberations is insufficient. By providing a framework for the complex interactive relationships between various “deliberative activities”, the systemic approach enables us “to identify which standards to employ when assessing the deliberative performance of a system as a whole” (Elstub et al., 2016, p. 140; see also Boswell, Hendriks, & Ercan, 2016; Dryzek, 2009).

There are two exemplary approaches for the measurement of deliberative performance on the basis of the systemic approach that shall serve as a starting point for our elaborations: Ercan et al. (2017, p. 197) “offer an interpretive response to...the empirical questions posed by the systemic turn”. But even though they are able to identify the contribution of qualitative case studies to the study of deliberative systems, they are aware of the fact that “interpretive studies are typically limited to discrete or small-n case studies” (Ercan et al., 2017, p. 206). John S. Dryzek (2009) attempts to evaluate the “deliberative capacity” of deliberative systems by referring to criteria of authenticity⁴ of the respective processes, their inclusiveness and their consequentiality. In doing so, he takes the core theoretical ideas of scholars of the systemic approach for granted and translates them into operationalizations. Thus, his approach seems to be limited to scholars who (to a large extent) accept the systemic approach’s premises, which means that it is hardly compatible with the measuring approaches of numerous other scholars who base their work on other theoretical foundations.

To avoid such problems, we will use the following parts of this article to propose a *modular* approach that is compatible with different indices of democratic performance or quality. The four PMMDs and the decisions

that are to be made in these steps offer a large degree of flexibility since they can be adjusted according to the theoretical and empirical focus of the researcher. Furthermore, the approach makes it possible to specifically address these challenges of scaling up the measurement of micro deliberation in conceptualization, measurement, and aggregation (Munck & Verkuilen, 2002). We are aware of the fact that an empirical specification of the systemic approach faces fundamental problems.⁵ Nevertheless, we use the *conceptual framework* of the systemic approach (Boswell & Corbett, 2017; Dryzek, 2009, 2016a, 2016b; Mansbridge et al., 2012) as a point of departure for our suggestions, as it seems to be the only framework available so far that does justice to the importance of the interactive relationships between different forms of deliberation for the deliberative performance of the political system as a whole.

3. “Parameters for Measuring Democratic Deliberation” at the Systemic Level

3.1. Core Elements of the Systemic Approach

The systemic approach considers deliberation to be an “emergent property” (Niemeyer et al., 2015): the deliberative quality is a property of the system as a whole and cannot (1) be located in a specific part of the system (locus) or (2) be reduced to a mere aggregation of other qualities of the political system. In the different *loci*, deliberations of various degrees of formality take place. One of Mansbridge’s (1999) original concerns was to include “everyday political talk” in the analysis of the deliberative performance of political systems (see also Conover & Searing, 2005, pp. 269f.). But this does not mean that the significance and deliberative character of formal institutions (such as parliaments or courts) should be underestimated or “downplayed” (Gaus, 2016, p. 511). To evaluate the deliberative performance of a political system, one has to consider formal *and* non-formal, institutionalized *and* non-institutionalized practices as well as “the interdependence of sites within a larger system” and their interactions (Mansbridge et al., 2012, p. 1; see also Bohman, 2012, p. 73).

The framework proposed in this article relies on three major claims of the systemic approach:

1. Deliberations in *different loci* (formal institutions, informal political talk, everyday conversations, etc.) of the political system are relevant;
2. In these loci, we will find *more or less formalized deliberative practices*: assumedly, in a federal court the standards for “good deliberative performance” and adequate ways of “taking and giving of reasons” will be much higher than in less for-

⁴ “Authenticity can be understood in light of the tests just introduced (deliberation must induce reflection in noncoercive fashion, connect particular claims to more general principles, and exhibit reciprocity)” (Dryzek, 2009, p. 1382). Schouten, Leroy and Glasbergen (2012) operationalize this concept by using the criteria of the DQI (Steiner et al., 2004).

⁵ Boswell and Corbett (2017, p. 4) identify the major challenges of empirical applications of the systemic approach that result from a lack of empirical specification, such as the measurement of “the deliberative effects of non-deliberative acts”.

mal contexts such as the public sphere or “everyday talk” which might be considered relevant;

3. There are *interactive relationships* between these deliberations that need to be considered in the attempt to determine the overall deliberative performance of the political system: the deliberative quality of the whole democratic political system is not the same as the mere accumulation of the deliberative performances in the individual loci. Instead, the interactions between deliberations in these loci need to be taken into account, as well as the performance within the loci.

In the following, we aim to propose a measurement approach compatible with different theories of democracy and concepts of democratic deliberation by building on these core elements of the systemic approach. In order to offer a useful guideline for future research, we name the relevant conceptual decisions (section 3.2.) that need to be made in the process of conceptualization, operationalization, and aggregation of “democratic deliberation”.

3.2. Parameters for the Conceptualization and Operationalization of Democratic Deliberation (PMMDs)

In this section, we suggest four parameters that need to be considered when measuring democratic deliberation at the macro level. By adjusting them, this measure can be made compatible with different conceptual frameworks and measurement approaches. To show the relevance of these conceptual decisions, we present examples to show why and to what extent the adjustment of the respective parameter makes a difference to the measurement of democratic deliberation and the results of the measurement process.

3.2.1. First Parameter: Theory of Democracy

The aim of this article is to develop suggestions for the measurement of democratic deliberations that are compatible with different measures of democracy and democratic performance. Even though the most frequently cited measurement approaches seem to measure the very same object (democracy or democratic performance), they refer to different democratic theories. This applies in particular to the way democratic institutions and their interactions are described as well as the way legitimacy is thought to be generated. Accordingly, the first conceptual question to be asked is which theory of democracy is at the heart of the measurement approach adopted. This is actually not only a preliminary question as it has rather important implications for the following steps.

Obviously, there is such an extensive range of theories of democracy that any attempt to summarize them here would be pointless.⁶ Instead of offering a com-

prehensive account of all (possible) choices that might need to be considered regarding the theory of democracy adopted, we want to illustrate the importance of the careful adjustment of parameter one with the help of one major division between contemporary schools of democratic thought: liberal and deliberative democratic theory. What are the implications of choosing one or the other, and to what extent does an affiliation with either side impact the decisions made in the process of measuring democratic deliberation? The fundamental difference between liberal and deliberative theories’ understanding of democracy is the logic of democratic legitimacy that is presupposed: while liberal theories follow the premise that an adequate representation of pre-political or endogenous preferences is the core criterion, deliberative theories assume that preferences are exogenous and legitimacy is generated by an inclusive debate that ideally results in consensus (or at least compromise). Thus, the role ascribed to democratic deliberation differs in both models. In the case of a liberal theory, its function is to validate endogenous preferences and thereby improve the epistemic quality of decisions reached in debates in representative institutions or to justify elite decisions vis-à-vis the public. In a deliberative theory, the democratic criterion is that the addressees of law also need to be the authors of law. This is the result of an inclusive deliberative process, which ideally leads to a consensual agreement between all stakeholders.

Depending on the theoretical framework adopted, there can be path dependencies for the measurement of democratic deliberation. First, the concepts of deliberation implied by the theory are likely to differ: a liberal model might value “(good) deliberation” mostly for its epistemic benefits and define it accordingly, whereas a deliberative understanding is much more likely to presuppose a procedural concept of deliberation (see *parameter two*, section 3.2.2.). The specific definition of “good deliberation” has important implications for the evaluation of the deliberative performance of democratic political systems: the operationalization, measurement, and even the loci considered to be relevant for the total score in this dimension (see *parameter three*, section 3.2.3.) will depend on the adjustment of this first parameter.

3.2.2. Second Parameter: Concept of Deliberation

The second parameter—a decision for a clearly defined concept of deliberation—is of particular importance, as there is an intense theoretical debate on what deliberation “really” is. This theoretical debate has serious implications for empirical research on democratic deliberation, as:

[this] lack of agreement about what constitutes deliberation makes it extremely difficult for empirical

⁶ *Varieties of Democracy*, for example, considers seven theoretical approaches of democracy and derives the respective “principles” from the thinking underlying them (Coppedge et al., 2016, pp. 4–6).

researchers to address the claims of normative theory. How can one safely assert that deliberation has occurred when there are no necessary and sufficient conditions routinely applied to this concept? (Mutz, 2008, p. 526)

Obviously, we cannot offer a comprehensive discussion of all concepts of deliberation here, but we want to point out two particularly important decisions that have to be made in this step. Before that, there are two preliminary questions scholars should ask: do their general measurement approach and the theoretical understanding presupposed by this approach (see *parameter one*) imply a definite and non-ambiguous definition of deliberation? And if so, should this concept of deliberation be used in the measurement of democratic deliberation as well? If both questions are answered negatively, the scholar needs to decide upon a concept of deliberation. In this step, two questions are of particular importance: (1) *is the criterion for “high deliberative performance” an epistemic or a procedural criterion?*⁷ (2) *Do I want to adopt a wide or a narrow concept of deliberation?* Both questions will be elaborated on in this section.

Question (1). While a procedural concept of deliberation would evaluate the deliberative performance of a system by the characteristics of the process(es) of deliberation (inclusiveness, fairness, etc.), an epistemic concept would evaluate the performance based on the output of this very process and the conformity of this outcome to an external criterion of “rightness” (Estlund, 2008). From a strictly theoretical perspective epistemic and procedural deliberation seem to be incompatible, therefore an explicit decision for one of them would be necessary. Nevertheless, this theoretical issue is not as pressing in a more pragmatic (empirical) approach. Deliberation in actual political systems can obviously have different functions simultaneously; it can promote the epistemic value of the decisions and the inclusion of all people affected by it, and it can, of course, be valued for both.⁸ In evaluating deliberative performance, one nevertheless needs to be aware of the fact that the scores for achieving the epistemic and the procedural goal can vary independently: expert deliberation can be highly beneficial for generating a “qualified” decision by being exclusive at the same time. In this article, we do not want to argue in favor of one concept or another, but simply want to raise awareness of the fact that different “kinds” of deliberation (depending on theoretical assumptions and chosen loci) might need to be evaluated by different standards.

Question (2). There is a broad range of concepts of deliberation applying more or less rigid standards. In the orig-

inal normative approaches, a narrow concept of deliberation was used: “deliberation” meant the taking and giving of reasons in the strictest sense, i.e., the exchange of rational (non-emotional), neutral, impartial arguments (Cohen, 1989; Habermas, 1996). In the confrontation with diversity theories, “we see a definite expansion of the sorts of things that could be considered arguments and reasons” (Chambers, 2003, p. 322). Partly as a result of “deep theorizing about reason”, and partly as a “result of confrontations with real-world practices”, a stretching of the original concept took place by taking into account that there are actually different “styles” and “cultures” of communication and reason-giving (Chambers, 2003).

Coming from the framework of the systemic approach, it seems reasonable to lower the standards for what counts as deliberation (as taking and giving of “good” reasons) in specific contexts. For example, “average citizens have few opportunities to deliberate rigorously in formal institutional settings. Most of their political discussions are therefore quite unstructured” (Conover & Searing, 2005, pp. 269f.).

If we regard different institutional settings, we also need to consider that these different deliberations should be “evaluated...by different standards” (Christiano, 2012, p. 28): an evaluation of the deliberative performance of a federal court and everyday political talk using the same standards would hardly make sense (cf. Christiano, 2012).

Although it is “a core axiom of the deliberative systemic approach: that non-deliberative practices can have positive systemic deliberative consequences, and as such should be treated as part of the system” (Dryzek, 2016a, p. 211), we object to any attempt to stretch the concept of deliberation too far. The inclusion of *non-deliberative practices* in the measurement of the overall deliberative performance of a political system has been criticized by various scholars (prominently: Owen & Smith, 2015, who suggest a *reductio ad absurdum* of this claim; see Dryzek, 2016b, p. 12).⁹ Additionally, an extensive lowering of standards would miss the point of setting a *normative standard* for the evaluation of the performance and quality of democracies: “A too realistic ideal is merely an apology for the status quo” (Neblo, 2007, p. 536; see also Elstub et al., 2016, p. 146). This does not mean that it is not possible to assume “a continuum of deliberative standards for assessing the parts of the system”, but only that they need “to be kept normatively robust and stringent” (Elstub et al., 2016, p. 146). Thus, on the one hand, a feasible and realistic approach (Bohman, 1998) that is compatible with the systemic approach cannot presuppose only “rational, reasonable, etc.” exchanges of logically valid arguments and “good” reasons in the strictest sense. On the other hand, if the concept is stretched too

⁷ In the theoretical debate, the controversy between “proceduralists” and “epistemic democrats” is one of the major cleavages (Estlund, 2008; Peter, 2007, 2013; for a discussion of this controversy see Fleuß, 2017, Chapter 3).

⁸ In the systemic approach, deliberation can fulfil three functions (epistemic, ethical, democratic) in different degrees. These functions partly correspond to different but not incompatible theoretical understandings of democratic deliberation (Mansbridge et al., 2012).

⁹ The validity of this *reductio* would have to be discussed in detail. Nevertheless, it highlights a problematic issue about the inclusion of *non-deliberative practices* in the evaluation of the *deliberative* performance of democracies.

far (see Steiner, 2008), there is no normative standard to compare deliberative performance with at all.

Therefore, we will stick to the claim that “deliberation” implies at least that the exchange of arguments and reasons *of some kind* occurs and we suggest that bargaining or story-telling should not be considered as “deliberative practices” (cf. Bächtiger & Wyss, 2013). Although this means that we exclude certain “communicative styles” from the concept of deliberation, there still remains a range of concepts of deliberation with a variety of scopes that might refer to a different range of phenomena, which (depending on the researcher’s conceptual decision) would have to be measured.

3.2.3. Third Parameter: Loci of Deliberation

In this section, we present a systematized list of loci that is suitable for the comparative measurement of democratic deliberation. Following Conover and Searing (2005, p. 270), we assume that deliberations relevant to assess the deliberative performance of the whole system can take place¹⁰ in three arenas of decreasing degree of formality:

(1) *Highly formal deliberations* “occur within institutions such as national courts, parliaments, and civil science departments” (Conover & Searing, 2005). These deliberations are probably most compliant with high standards of rationality;

(2) *Semi-formal deliberations* are “conversations between constituents and government officials, and conversations in political parties, interest groups, and the media” (Conover & Searing, 2005). Here, a lowering of the “rationality-standards” is probably necessary for the evaluation of the deliberative performance in these spheres;

(3) *Informal deliberations* are the “less deliberative everyday discussions among political activists, attentive publics and general publics; a form of political talk that is essential to the system’s democratic character” (Conover & Searing, 2005). We expect informal deliberations to be least compliant with demanding normative standards of rationality and impartiality.

Depending on theoretical and conceptual decisions, different potential loci of deliberation will be selected and prioritized for the evaluation of the overall deliberative performance of a political system. In the selection of these loci, the parameters one and two and the choices made in these steps are relevant as well: a liberal theory, for example, will suggest a different relative weight of parliamentary deliberation than a deliberative theory

and might not consider some deliberations named in category (3) to be relevant for the deliberative performance of the political system at all. Also, the selection of a wide or narrow concept of deliberation will have an impact on the selection of the loci: depending on how far one is willing to “stretch the concept”, a different range of phenomena will be included in the measurement conducted on this basis.

From categories (1)–(3), we can derive a systematized list of potential loci of deliberation, which offers a much more useful framework for an empirical analysis than the enumerations given by Mansbridge et al. (2012, pp. 2, 7, 10; see also Conover & Searing, 2005; Erkan et al., 2017). This procedure also matches the *loci* to spheres in which different *kinds of deliberative practices* (which have to be evaluated by different standards) take place.¹¹ As we are about to demonstrate, the loci of each of the corresponding categories (1)–(3) require a different measurement approach, depending on the nature of the deliberative practices in question, and—pragmatically speaking—the accessibility of data. This is why, after the explication of each category, we will point out what has to be considered if a measurement of deliberations occurring in the respective loci is envisaged. Depending on what kind of deliberation is to be measured and what kind of data is available for the respective deliberation, different methodological approaches need to be considered.

The logical assignment of the different loci to the categories proposed by scholars has so far been somewhat vague and rarely more precise than: outlining “a spectrum of venues for deliberation, including representative assemblies, public assemblies, the public sphere, and everyday talk, and ‘moving along this range entails moving along a similar range, from formal to informal’” (Elstub et al., 2016, p. 145). Not all potential loci of deliberation can be assigned to just one category: different kinds of deliberation can appear in one locus, though usually there is a tendency for certain kinds of deliberation to occur in a certain locus.

With regard to measurement approaches and operationalizations, the loci in the different categories need to be treated quite differently. This is partly determined by the availability of data on deliberations—while parliamentary deliberations are generally recorded, deliberation in less formalized loci such as marketplaces usually happens spontaneously, without audience or record. Furthermore, different kinds of deliberation occur within different frameworks in terms of timeframe, the presence of the participants, and strictness of rules. On one hand, there are the contributions of MPs to parliamentary debates which usually follow a general pattern, have a certain timeframe (according to the respective protocol), and are restricted to defined topics and a

¹⁰ In the following, we refer to “potential loci of deliberation” for two reasons: First, comparative research faces the problem that depending on the institutional reification of the “deliberative system” in question, certain loci might not exist. Second, the existence of a specific *locus* does not necessarily mean that (relevant) deliberation is actually taking place in this space.

¹¹ Thereby we attempt to provide a framework that is more systematic and at the same time (due to its modular character) more flexible than previous research using the deliberative systemic approach.

defined type of language. On the other hand, there are online debates whose participants are generally free in their expressions concerning structure and language, as well as in their use of pictures, videos or other sign systems such as hashtags, likes and emoji. Online debates do not have limitations in terms of time or space; anyone can log in from anywhere in the world anytime. To analyze these different modes of deliberation scholars of deliberative performance need to use different methods. We will give examples for each category of deliberation in order to illustrate what measurement approaches can be used.

Loci of “highly formal deliberations” (1). As cited above, Conover and Searing (2005, p. 270) assign loci such as national courts, parliaments, and civil science departments to the level of highly formal deliberation (in their terminology: “structured deliberation”). However, we propose to include in this category only loci that are constitutionally or otherwise legally installed, that follow certain (procedural) rules while deliberating and that have the power to make collectively binding decisions.¹² Thus, we differ from Conover and Searing by excluding any locus of deliberation that does not meet those criteria (such as the civil science departments they suggest). In addition, “highly formal deliberation” is not necessarily restricted to deliberation taking place in constitutional or representational bodies: certain “democratic innovations” (such as mini publics) can be subsumed under this category if they are empowered to make collectively binding decisions (Fung, 2006). To measure the deliberative performance in loci of highly formal deliberation, scholars can use minutes and reports of the deliberations, as well as written statements or legislative proposals prepared in advance of the deliberations.

Loci of “semi-formal deliberations” (2). Conover and Searing (2005, p. 270) define semi-formal deliberations as “conversations between constituents and government officials, and conversations in political parties, interest groups, and the media”. This correlates with what Habermas calls “the public sphere”, which is situated around the political center and which functions as the transition sphere of political ideas and arguments to that center. Habermas (1996) also includes journals, interest groups, clubs, professional associations, academies and universities, as well as grass root initiatives. We would like to complement this list with NGO-related spaces and meetings, trade unions, and other lobby groups. Thus, this category remains quite vague and cannot be described by more specific criteria than: (1) it is *the zone where members of the political elite and members of the public sphere deliberate, or where such encounters*

are prepared, and (2) there is a *certain degree of institutionalization*. Again, we differ from Conover and Searing who assign party deliberations to this category. Measuring the deliberative performance in these loci can be attempted with the help of minutes (if existent) or by interviewing insiders and experts.

Loci of “informal deliberations” (3).¹³ In contrast to the other two categories, informal deliberation is not at all institutionalized (in the sense of being regulated by formalized rules), i.e. the “less deliberative everyday discussion among political activists, attentive publics and general publics” (Conover & Searing, 2005, p. 270). Loci of this kind of deliberation can be “ad hoc forums, or online spaces within which ordinary citizens, members of social movements, and civil society actors can engage in discussion and debate” (Smith, 2016, p. 154), offline and online comments in response to news items, as well as marketplaces and their culturally specific equivalents. Sources of data for measuring deliberative performance can again be interviews with insiders and experts. Furthermore, online deliberation within social media platforms and in comment feeds are especially helpful in that they enable scholars to explore informal deliberation in great detail with the help of computational text mining devices.

3.2.4. Fourth Parameter: Aggregation Rule

As previously stated (section 2), the deliberative quality of the entire democratic political system does not equal the mere accumulation of the deliberative performances in the individual loci. Rather, the interactions between deliberations in these loci also need to be taken into consideration. So far, the interactive relationships between deliberations in different loci have been addressed in case studies (Boswell et al., 2016; Ercan et al., 2017) and in various approaches comparing deliberative systems (Boswell & Corbett, 2017). However, a comprehensive approach to taking the interactive relationships between different loci systematically into account, instead of merely scaling up micro level measurement of deliberative performance, is still missing. The fourth parameter addresses questions and choices that should be considered when developing such an aggregation rule.

In line with the systemic approach, we regard two kinds of interaction to be crucial for the evaluation of the deliberative performance of political systems at the macro level: the *transmissions* between deliberative procedures taking place in the more or less formalized spheres as well as their (potential) *complementarity*. Thus, these two should be reflected in the aggregation rule. Generally, aggregation rules consist of three

¹² Depending on their power and authority, some constitutional courts fall into this category with regard to some judicial matters. Since constitutional courts are not conventionally regarded as part of the political system, we exclude them from our further discussion.

¹³ One important innovation of the systemic approach was the inclusion of non-deliberative practices in the evaluation of the overall deliberative performance. In the exposition of parameter (2), we explained why *explicitly* non-deliberative practices are to be excluded in the measurement. In addition to the normative and conceptual considerations regarding this matter outlined above, there are pragmatic reasons for excluding certain communicative styles from the analysis (see Boswell & Corbett, 2017, p. 15).

kinds of element: variables, weights, and operations. In our framework, the variables describe the deliberative performances measured for the different loci (see Parameter 3). In the following, we will show that the weights can be based on the degree of transmission and that the operations depend on the relationships between the loci as well as on their complementarity.

3.3. “*Transmissions*” and *Weighting*

The aggregation rule needs to take into account that the results of deliberative processes reached in different loci and “spheres” (1–3) “must be proliferated across and among sites so that they can be challenged and ‘laundered’ through the system” (Boswell et al., 2016, p. 264). There have already been some attempts at capturing this “interplay” of deliberations, which is crucial for the deliberative performance of the whole system (Boswell et al., 2016). However, a systematic way that is compatible with different indices of democracy is a serious challenge.

Mansbridge et al. (2012, p. 23) do not offer an explicit definition of the interactions between loci, but rather speak (sometimes in a metaphorical way) of “coupling” (see also Hendriks, 2016a, p. 44). While the concept of coupling focuses on the relationships between the loci of the deliberative system (tight coupling of loci vs. loose coupling of loci) (Hendriks, 2016a), the concept of transmissions refers to the transfer of reasons given and results achieved in deliberations among the various loci. For feasibility reasons, we suggest that scholars use the concept of transmissions for the measurement of deliberative performance: the identification of reasons and results of reason-giving processes that might or might not be transferred to another locus seems to be much easier than the measurement of the degree of “coupling” of the loci of the respective deliberative practices.

There are three ways of tracking the transmissions of topics between different loci and assigning them a score in order to compare different democracies in terms of transmissions. Firstly, scholars could track certain topics as they evolve throughout the system (as done by Ercan et al., 2017, pp. 201-203), counting the number of loci they pass through as well as recording whether they have been present in all three categories of deliberation. Secondly, scholars could track certain individuals who potentially transmit ideas from one locus to another (cf. Mendonça, 2016) in terms of how many different loci in which categories they frequent and how many topics they pass from one to another.

However, we strongly recommend a third approach: observing certain loci with regard to where the transmitted elements (that is, deliberated ideas, reasons, resolutions, etc.) come from and where they are transmitted to (as done by Boswell et al., 2016, pp. 270-273; Hendriks, 2016a). Translated into the quantitative measurement of deliberative quality that would be: how many elements are transmitted? And how many other loci are involved

in these transmissions? That approach is the most feasible for three reasons. It (1) is easily integrated into the measurement of the loci, since these loci are being assessed anyway. Thus, scholars could use elaborate methods like participant observation, but they could also take the materials they already use for assessing the deliberative quality and browse them with the help of computers for citations, expert opinions, and references to news articles, activist groups and such. Since that would only deliver a fairly accurate approximation of transmission, it should be complemented with the tracking of certain topics (cf. the first approach) in order to at least gauge the accuracy of the approximation. Furthermore, this approach would (2) be far more systematic since all loci in the study would be included and could be assessed by the same methods and it would (3) provide an approach for the weighting within an aggregation rule. The scope of transmission of each locus could be used as a weight, either in an inclusive sense for the whole system, or for each respective locus. The theoretical implication in terms of the deliberative systemic approach would be: the larger the scope of transmission, the better the deliberative quality. Furthermore, the importance of the deliberation in one locus could be assessed with that method as well: the more it is referred to (and refers to itself), the more important it becomes for the whole system—thus providing another option for a systematic aggregation rule for the measurement of macro deliberation. Consequently, for reasons of feasibility and compatibility with different democracy indices that approach should be the most suitable for most studies of deliberative performance. However, the choice of method always depends on the aim of the study and the instrument it is to be integrated within.

3.4. “*Complementarity*” and *Operations*

One fundamental assumption of the systemic approach is that “[t]hough there may be little or no perfect democratic deliberation in any site, the collective work done across the system may still produce a suitably deliberative democratic whole” (Boswell et al., 2016, p. 263). Accordingly, an aggregation rule taking this line of thinking seriously needs to take account of the complementarity of deliberation in different loci, and therefore the substitutability of the deliberation in one locus. There are several assumptions to be made about what the defining criteria for the degree to which a locus is substitutable are. Firstly, it might depend on the level of formalization of the locus. Secondly, it might depend on the (legally ascribed) political importance of that locus, which correlates with—but is not identical to—the degree of formalization. Thirdly, it might depend on its importance for the deliberative system, which can be assessed by the method recommended above—the more transmission links and transmitted elements, the higher the importance. Fourthly, it might depend on the structure of the locus. On that line of thought, Boswell and Corbett

(2017) present an approach to compare deliberative systems via “family resemblances”:

At its centre are recurring “traits” that come and go, to varying degrees, across units within the same broad family. Such traits might include institutional variants, but they tend to entail a decentred, interpretive account of these institutions—one that sees them not as given, but as constructed and continually reproduced through social interaction.

This approach can be transferred to the level of loci. “Traits” can integrate some of the criteria mentioned above (such as the level of formalization). Loci, with similar structural traits (thus belonging to one “family”), could be deemed complementary, and the more members of the respective family, the more substitutable the single locus. However, some loci might not be substitutable at all, in spite of family resemblances to other loci. Examples are deliberations in parliaments and courts. That “unsubstitutability” has to be marked as a trait as well.

Although some democracy indices’ aggregation rules seem very elaborate,¹⁴ there are two mathematical operations which all aggregation rules are based on: addition and multiplication. Those rules imply different theoretical assumptions concerning the relationship of the attributes that are to be aggregated: “If one’s theory indicates that both attributes are necessary features, one could multiply both scores, and if one’s theory indicates that both attributes are sufficient features, one could take the score of the highest attribute” (Munck & Verkuilen, 2002, p. 24), or, alternatively, cumulate all scores. Consequently, complementary deliberations in loci of one family could be added up to one score, while deliberations in non-complementary loci should be multiplied. In the first case, there are two options. Either, the total scores of deliberative performance in each locus are cumulated, or the scores for the chosen criteria for deliberative performance (Parameter 2) are cumulated across loci, prior to using the chosen aggregation rule for deliberative performance—thus, the deliberation within one family of loci would be treated as one truly complementary unit. In the second case, low deliberation scores in “unsubstitutable” loci would vastly lower the total score, and a zero would reduce the overall score to nil.

The choices to be made concerning the aggregation rules depend on the selection of the democratic theory on one hand, and on the understanding of deliberation on the other. For example, an index based on liberal democratic theory will probably place greater weight on highly formal deliberation (by individual weights as well as the use of multiplications) than an index based on deliberative democratic theory. Thus, the fourth parameter, again, builds upon the choices made concerning the previous parameters.

¹⁴ For example, V-Dem uses factor indices that are based on point estimates drawn from a Bayesian factor analysis model (cf. Coppedge et al., 2016, p. 11).

4. Conclusion: Challenges of Measuring Democratic Deliberation

In this article, we argued for the need to include the measurement of democratic deliberation into the evaluation of the democratic performance of political systems. Accordingly, we developed guidelines for a theoretically grounded measurement of deliberative performance at the macro level. Since we intended to make our suggestions compatible with different available approaches to measuring democratic performance, we proposed a modular approach. The core elements of this approach are the four PMMD that can be adjusted in various ways to fit in with the measurement approach adopted. The specific indicators which should be used to conduct the measurement have to be decided upon in accordance with the specific adjustments of these PMMDs. In the suggestions we provided concerning these parameters, we tried to do justice to the specific requirements of the measurement of deliberation at the macro level which are to a large extent based on the systemic approach (Beste, 2016; Dryzek, 2015, 2016a, 2016b; Mansbridge et al., 2012).

We are aware of the fact that this specific theoretical framework not only has its own theoretical pitfalls (Hendriks, 2016b; Owen & Smith, 2015) but that it also carries intricate methodological challenges, especially in terms of feasibility. The most complicated challenge is probably how to adequately reflect the interactive relationships of deliberations in different loci—their transmission and their complementarity—in the aggregation rule. Here, future research should further address not only the question of how transmissions or complementarity can be adequately theorized, but also how they can be measured in practice at the macro level in comparative large-n studies.

We firmly believe that it is impossible to develop a one-size-fits-all solution for this issue. Rather, the solution adopted for the integration of the measurement of deliberative performance will to a large extent be dependent on the theoretical, conceptual, and methodological “parameters” previously chosen by the researcher.

Conflict of Interests

The authors declare no conflict of interests.

References

- Bächtiger, A., & Wyss, D. (2013). Empirische Deliberationsforschung—eine systematische Übersicht [Empirical deliberation—A systematic review]. *Zeitschrift für Vergleichende Politikwissenschaft*, 7(2), 155–181.
- Beste, S. (2016). Legislative frame representation: Towards an empirical account of the deliberative systems approach. *Journal of Legislative Studies*, 22(3), 295–328.

- Bohman, J. (1998). Survey article: The coming of age of deliberative democracy. *Journal of Political Philosophy*, 6(4), 400–425.
- Bohman, J. (2012). Representation in the deliberative system. In J. Parkinson & J. Mansbridge (Eds.), *Deliberative systems: Deliberative democracy at the large scale* (pp. 72–94). Cambridge: Cambridge University Press.
- Boswell, J., & Corbett, J. (2017). Why and how to compare deliberative systems. *European Journal of Political Research*, 56(4), 801–819.
- Boswell, J., Hendriks, C. M., & Ercan, S. A. (2016). Message received? *Critical Policy Studies*, 10(3), 263–283.
- Chambers, S. (2003). Deliberative democratic theory. *Annual Review of Political Science*, 6(1), 307–326.
- Chambers, S. (2012). Deliberation and mass democracy. In J. Parkinson & J. Mansbridge (Eds.), *Deliberative systems: Deliberative democracy at the large scale* (pp. 52–71). Cambridge: Cambridge University Press.
- Cohen, J. (1989). Deliberation and democratic legitimacy. In A. Hamlin & P. Petit (Eds.), *The good polity. Normative analysis of the state* (pp. 17–34). Oxford: Polity Press.
- Cohen, J. (1997). Procedure and substance in deliberative democracy. In J. Bohmann & W. Rehg (Eds.), *Deliberative democracy. Essays on reason and politics* (pp. 87–106). Oxford: Blackwell.
- Christiano, T. (2012). Rational deliberation among experts and citizens. In J. Parkinson & J. Mansbridge (Eds.), *Deliberative systems: Deliberative democracy at the large scale* (pp. 27–51). Cambridge: Cambridge University Press.
- Conover, P. J., & Searing, D. D. (2005). Studying ‘everyday political talk’ in the deliberative system. *Acta Politica*, 40(3), 269–283.
- Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S. E., Teorell, J., Altman, D., . . . Zimmerman, B. (2016). *V-Dem codebook v6*. University of Gothenburg: Varieties of Democracy (V-Dem) Project.
- Dryzek, J. S. (2009). Democratization as deliberative capacity building. *Comparative Political Studies*, 42(11), 1379–1402.
- Dryzek, J. S. (2010). *Foundations and frontiers of deliberative governance*. Oxford: Oxford University Press.
- Dryzek, J. S. (2015). Deliberative engagement: The forum in the system. *Journal of Environmental Studies and Sciences*, 5(4), 750–754.
- Dryzek, J. S. (2016a). Symposium commentary: Reflections on the theory of deliberative systems. *Critical Policy Studies*, 10(2), 209–215.
- Dryzek, J. S. (2016b). The forum, the system, and the polity. *Political Theory*, 44(1), 1–27.
- Elstub, S. (2015). A genealogy of deliberative democracy. *Democratic Theory*, 2(1), 100–117.
- Elstub, S., Ercan, S., & Mendonça, R. F. (2016). Editorial introduction: The fourth generation of deliberative democracy. *Critical Policy Studies*, 10(2), 139–151.
- Ercan, S. A., Hendriks, C. M., & Boswell, J. (2017). Studying public deliberation after the systemic turn: The crucial role for interpretive research. *Policy and Politics*, 45(2), 195–212.
- Erman, E. (2016). Representation, equality, and inclusion in deliberative systems: Desiderata for a good account. *Critical Review of International Social and Political Philosophy*, 19(3), 263–282.
- Estlund, D. M. (2008). *Democratic authority. A philosophical framework*. Princeton, NJ: Princeton University Press.
- Felicetti, A., Niemeyer, S., & Curato, N. (2016). Improving deliberative participation: Connecting mini-publics to deliberative systems. *European Political Science Review*, 8(3), 427–448.
- Fishkin, J. (1995). Bringing deliberation to democracy: The British experiment. *The Good Society*, 5(3), 45–49.
- Fleuß, D. (2017). *Prozeduren, Rechte, Demokratie. Das legitimatorische Potential von Verfahren für politische Systeme [The normative legitimacy of democracies. On the limits of proceduralism]* (Doctoral dissertation). University of Heidelberg, Heidelberg, Germany. Retrieved from <http://archiv.ub.uni-heidelberg.de/volltextserver/23203/1/Fleuss-Prozeduren%20Rechte%20Demokratie-2017.pdf>
- Fung, A. (2006). Varieties of participation in complex governance. *Public Administration Review*, 66, 66–75.
- Gaus, D. (2016). Discourse theory’s sociological claim: Reconstructing the epistemic meaning of democracy as a deliberative system. *Philosophy & Social Criticism*, 42(6), 503–525.
- Geißel, B., & Newton, K. (Eds.). (2012). *Evaluating democratic innovations*. New York, NY: Routledge.
- Habermas, J. (1996). *Between facts and norms: Contributions to a discourse theory of law and democracy*. Cambridge, MA: MIT Press.
- Hendriks, C. M. (2016a). Coupling citizens and elites in deliberative systems: The role of institutional design. *European Journal of Political Research*, 55(1), 43–60.
- Hendriks, C. M. (2016b). Integrated deliberation: Reconciling civil society’s dual role in deliberative democracy. *Political Studies*, 54(3), 486–508.
- Kaase, M. (1984). The challenge of the ‘participatory revolution’ in pluralist democracies. *International Political Science Review*, 5(3), 299–318.
- Mansbridge, J. (1999). Everyday talk in the deliberative system. In S. Macedo (Ed.), *Deliberative politics. Essays on democracy and disagreement* (pp. 211–239). Oxford: Oxford University Press.
- Mansbridge, J. (2007). Deliberative democracy or democratic deliberation? In S. Rosenberg (Ed.), *Deliberation, participation and democracy* (pp. 251–271). London: Palgrave Macmillan.
- Mansbridge, J. (2010). Deliberative polling as the gold standard. *The Good Society*, 19(1), 55–62.
- Mansbridge, J., Bohman, J., Chambers, S., Christiano, T., Fung, A., Parkinson, J., . . . Warren, M. E. (2012). A systemic approach to deliberative democracy. In J.

- Parkinson & J. Mansbridge (Eds.), *Deliberative systems* (pp. 1–26). Cambridge, MA: Cambridge University Press.
- Mendonça, R. F. (2016). Mitigating systemic dangers: The role of connectivity inducers in a deliberative system. *Critical Policy Studies*, 10(2), 171–190.
- Merkel, W., Bochsler, D., Bousbah, K., Bühlmann, M., Giebler, H., Hänni, M., . . . Wessels, B. (2016). *Democracy barometer. Methodology* (Version 5). Aarau: Zentrum für Demokratie.
- Munck, G. L., & Verkuilen, J. (2002). Conceptualizing and measuring democracy: Evaluating alternative indices. *Comparative Political Studies*, 35(1), 5–34.
- Mutz, D. C. (2008). Is deliberative democracy a falsifiable theory? *Annual Review of Political Science*, 11(1), 521–538.
- Neblo, M. A. (2007). Family disputes: Diversity in defining and measuring deliberation. *Swiss Political Science Review*, 13(4), 527–557.
- Niemeyer, S. (2014). Scaling up deliberation to mass publics: Harnessing mini-publics in a deliberative system. In K. Grönlund, A. Bächtiger, & M. Setälä (Eds.), *Deliberative mini-publics. Involving citizens in the democratic process* (pp. 177–202). Colchester: ECPR Press.
- Niemeyer, S., Curato, N., & Bächtiger, A. (2015). *Assessing the deliberative capacity of democratic polities and the factors that contribute to it*. Paper presented at the conference “Democracy: A Citizen Perspective”, Turku, Finland. Retrieved from [http://www.abo.fi/fakultet/media/33801/niemeyerbachtigercurato_](http://www.abo.fi/fakultet/media/33801/niemeyerbachtigercurato_assessingthecapacityofdeliberativesystems.pdf)
- [assessingthecapacityofdeliberativesystems.pdf](http://www.abo.fi/fakultet/media/33801/niemeyerbachtigercurato_assessingthecapacityofdeliberativesystems.pdf)
- O’Connor, T., & Wong, H. Y. (2015). Emergent properties. *Stanford Encyclopedia of Philosophy*. Retrieved from <https://plato.stanford.edu/entries/properties-emergent/#OntEme>
- Owen, D., & Smith, G. (2015). Survey article: Deliberation, democracy, and the systemic turn. *Journal of Political Philosophy*, 23(2), 213–234.
- Peter, F. (2007). Democratic legitimacy and proceduralist social epistemology. *Politics, Philosophy & Economics*, 6(3), 329–353.
- Peter, F. (2013). The procedural epistemic value of deliberation. *Synthese*, 190(7), 1253–1266.
- Schouten, G., Leroy, P., & Glasbergen, P. (2012). On the deliberative capacity of private multi-stakeholder governance: The roundtables on responsible soy and sustainable palm oil. *Ecological Economics*, 83, 42–50.
- Smith, W. (2016). The boundaries of a deliberative system: The case of disruptive protest. *Critical Policy Studies*, 10(2), 152–170.
- Steiner, J. (2008). Concept stretching: The case of deliberation. *European Political Science*, 7(2), 186–190.
- Steiner, J., Bächtiger, A., Spörndli, M., & Steenbergen, M. R. (2004). *Deliberative politics in action. Analyzing parliamentary discourse*. Cambridge: Cambridge University Press.
- Stevenson, H., & Dryzek, J. S. (2012). The discursive democratization of global climate governance. *Environmental Politics*, 21(2), 189–210.

About the Authors



Dannica Fleuß is a postdoctoral researcher at Helmut-Schmidt-University, Hamburg. She completed her PhD on proceduralist theories of democracy and is currently working on applications of deliberative theories as well as the corresponding measurement approaches. Her research interests include political philosophy, measuring democracy, modern theories of democracy, and political-cultural studies.



Karoline Helbig is a doctoral researcher for Political Theory at the Helmut-Schmidt-University, Hamburg. She studied sociology and mathematics and currently is working on her PhD project regarding the influence of digital algorithms such as search engines and news feeds on democratic societies, with a particular focus on their deliberative procedures. Her further research interests include measuring democracy, methods of social research, democratic theory, digitalization, and the sociology of knowledge.



Gary S. Schaal is full professor of Political Theory at the Helmut-Schmidt-University, Hamburg. He worked on a wide variety of topics, including contemporary normative political theory, emotions in the political sphere and digital democracy.