
Measuring popular and judicial deliberation: A critical comparison

Donald Bello Hutt*

This article compares instruments designed to measure deliberation in judicial and non-judicial settings. I thus provide a critical examination of different mechanisms deliberative democrats have designed to test what transpires when individuals deliberate from the perspective of ordinary citizens vis-à-vis the point of view of judges. From this appraisal, I conclude, first, that an examination of the literature on deliberation measurement brings to light several problems in the process of translating ideal deliberative theory into empirical evaluative schemes. Second, by relying on a critical examination of Conrado Hübner Mendes's work on deliberation in constitutional courts, I argue that those difficulties become starker when we try to assess the quality of judicial deliberation, given that our access to the courtroom is limited by the very structure of judicial procedures. Third, I argue that these two problems combined entail that idealizations of the courtroom as the forum in which ideal aspects of deliberative democracy are instantiated, are misguided, and should be avoided.

1. Introduction

This article compares instruments designed to measure deliberation in non-judicial settings, on the one hand, and deliberation in the courtroom, on the other. I survey different mechanisms designed by deliberative democrats to test what transpires when individuals deliberate from the perspective of ordinary citizens vis-à-vis the point of view of judges. From this appraisal, I conclude, first, that an examination of the literature on deliberation measurement brings into light a number of problems in the process of translating ideal deliberative theory into empirical evaluative schemes; second,

* FWO (Flanders Research Foundation) Postdoctoral Fellow, Institute of Philosophy KU-Leuven. This article was written during my time as a PhD candidate at the department of Political Economy, King's College London. I am grateful to the reviewers of this journal for helping me improve the manuscript. I also thank Adrian Blau, Lorenzo Zucca, Andreas Schäfer, Raphaël Kies, and Andrej Kristan for discussions and their helpful comments on earlier versions. Email: donaldemerson.bellohutt@kuleuven.be.

that in the case of courts, the structure of judicial procedures makes those difficulties deeper; and third, that these two problems combined indicate that idealizations of the courtroom as the forum in which ideal aspects of deliberative democracy are instantiated, are misguided, and should be avoided.

This approach matters for four reasons. First, the diversity of deliberative experiments and research on specific deliberative sites makes it impossible to cover all those instances in a comprehensive fashion. A more reasonable and manageable approach must focus on the instruments designed to evaluate those experiments. Second, although scholars have pointed to several problems faced by these measuring schemes, there are areas where there is still room for more criticism. Third, in spite of the staggering amount of research evaluating the deliberative performance of individuals in different settings, researchers have not measured how deliberation takes place in the courtroom. Hence, there is a gap to be filled. Fourth, because this gap does not square with the idea, endorsed by a number of prominent deliberative democrats, that courts are ideal *loci* for deliberation, these scholars mistakenly assume that the sort of dialogue and argument exchange taking place inside the courtroom instantiates an important number of ideal principles grounding deliberative democracy.

Deliberative democracy is a theory that, despite important substantive developments, is not currently in a position to fully test the adequacy of its practice to its principles. This does not mean that the project should be abandoned. Rather, it means that social scientists should continue working on fixing the defects and shortcomings of the tools they use to measure deliberation. It also means that there is little support for the claim that judicial procedures in particular ought to be seen as the exemplar channels of deliberation. Considering, on the one hand, the difficulties in measuring deliberative quality and quantity within sites that researchers can actually have access to, and considering their incapacity to check how and to what extent is deliberation happening inside the courtroom on the other, the notion that courts are ideal sites for deliberation becomes difficult to justify.

This article thus provides a novel explanation for social scientists' lack of attention on the measurement of judicial deliberation, and an original argument for the suggestion that if one prices deliberation, courts are not the best place to look for it. This is not because judges are inherently reluctant to debate but because their institutional environment and conditions prevent them and the parties at the trial from deliberating, and prevent researchers from checking how and whether judges deliberate among themselves, and with the parties at the trials conducted by them.

The article is structured as follows. Section 2 provides an overview of and makes critical remarks on the mechanisms designed to measure individuals' deliberative performance available in the literature, which I categorize as *popular*. I start by surveying the most discussed instruments that measure the quality and quantity of deliberation in these fora. In subsection 2.1, I then criticize these instruments on the grounds that they show deficiencies in their theoretical foundations, and in their understanding of and consideration for important aspects of ideal theory, such as participation, equality, and sincerity or truthfulness.

This part provides the context against which judicial deliberation is then analyzed in Section 3. It suggests that idealizations of any forum in terms of its deliberative capacity are to be avoided. Moreover, it gives way to the question addressed in Section 3 of why courts are not typically subjects of deliberation measurement.

Section 3 appraises deliberation measurement in the courtroom or, as I label it, *judicial* deliberation. In spite of the high regard that several prominent deliberative democrats show toward judges in relation to their deliberative capacities, and despite the efforts that social scientists and political theorists have put forth to investigate if, how, and to what degree individuals deliberate in such a varied group of settings, courts do not generally appear as a subject of examination in deliberation measurement studies. Hence, the empirical literature on the measurement of the quality of judicial deliberation is scant. Section 3 thus explores reasons for this neglect. It analyzes how the problem of deliberative measurement has been dealt with when it comes to courts, and shows that what we find in the literature is not empirical studies but ideal benchmarks. For this, I rely on Conrado Hübner Mendes's account of a deliberative constitutional court to show that, even in an ideal version, structural features of judicial procedures provide reasons to question that courts are ideal deliberative institutions. Overall, this part shows that there is not enough evidence to know how judges deliberate, let alone to claim that *judicial* deliberation is superior or inferior to *popular* deliberation.

Section 4 concludes that although there is much room for improvement on measurements of deliberation, the judicial side of the comparison shows deficiencies inherent to the structure of the procedures conducted inside the courtroom, especially, though not exclusively, because those discussions take place behind closed doors.

Overall, the article shows that there are two overarching reasons for not focusing on the judiciary when looking for ideal deliberative sites: on the one hand, the instruments available for assessing deliberative quality have non-negligible limitations. On the other, the ways in which judicial procedures are set prevent us from affirming that courts are deliberative institutions in the relevant sense demanded by theories of deliberative democracy.

2. Popular deliberation

This section gives a critical overview of the devices designed to measure deliberation in real-world fora. Given that the literature measuring *popular* deliberation is vast and extremely diverse, that it focuses on myriad issues, with different levels of salience, with higher or lower degrees of institutionalization, I will survey some of the most employed and discussed measuring instruments available. The diversity of the exercises of deliberation measurement makes it practically impossible to undertake a full survey. Yet, the instruments I now describe are among the most cited in the literature, and are representative of the most developed mechanisms designed to evaluate deliberation.

After this overview, I make critical remarks highlighting the limitations faced by these instruments. I do this with two objectives in mind. First, to show that the tools used by deliberative democrats to assess the quality and quantity of deliberation show non-negligible limitations in terms of their suitability to test the adequacy of deliberative practices to the principles underpinning deliberative democracy. Second, these critical remarks allow me to establish a background against which the claim that courts are ideal deliberative institutions can be later scrutinized in Section 3.

Empirical studies in deliberative democracy aim at determining whether individuals are able to engage in conversation with other citizens, mutually respecting each other, letting themselves be convinced by the unforced force of the better argument.¹ This is because even if deliberative democracy is generally seen as an ideal benchmark for decision-making processes, it is nonetheless rooted in actual practices and in the actual possibilities individuals have of engaging in debate and reason-giving.² The interplay between theory and the possibilities of its implementation has led researchers to think of ways to measure if and how much theory informs practice, and/or whether practice conforms to ideal theory.³

These efforts are, nonetheless, still work in progress and there is much room for improvement. Some mechanisms are limited to evaluate specific experiments, while others conceived of to have a more general application struggle to incorporate every possible element of ideal theory. Others have been superseded by more refined versions, which are in turn affected by other shortcomings. Moreover, there is no general agreement on working with a single scheme.⁴

¹ Georgia Warnke, *Communicative Rationality and Cultural Values*, in *THE CAMBRIDGE COMPANION TO HABERMAS* 120, 127 (Stephen K. White ed., 1995); JÜRGEN HABERMAS, *BETWEEN FACTS AND NORMS: CONTRIBUTIONS TO A DISCOURSE THEORY OF LAW AND DEMOCRACY*, §§ 305–306 (1996) [hereinafter HABERMAS, *BETWEEN FACTS AND NORMS*]; James Bohman, *Deliberative Democracy and Effective Social Freedom: Capabilities, Resources, and Opportunities*, in *DELIBERATIVE DEMOCRACY* 321, 322 (James Bohman & William Rehg eds., 1997); André Bächtiger et al., *Measuring Deliberation 2.0: Standards, Discourse Types, and Sequentialization*, ASH CENTER FOR DEMOCRATIC GOVERNANCE AND INNOVATION (Sept. 2009), available at https://ash.harvard.edu/files/ash/files/baechtiger_0.pdf; Valentin Gold, Katharina Holzinger, & Christian Rohrdantz, *Towards Automating the Measurement of Deliberative Communication*, 7TH GENERAL CONFERENCE OF THE EUROPEAN CONSORTIUM OF POLITICAL RESEARCH (ECPR), BORDEAUX, FRANCE (2013), available at http://kenbenoit.net/pdfs/NDATAD2013/Goldetal2013_Towards%20Automating%20the%20Measurement%20of%20Deliberative%20Communication.pdf (accessed Sept. 26, 2018); Jürgen Habermas, *Concluding Comments on Empirical Approaches to Deliberative Politics*, 40 ACTA POL. 384, 384 (2005) [hereinafter Habermas, *Concluding Comments*]; JOSÉ LUIS MARTÍ, *LA REPÚBLICA DELIBERATIVA: UNA TEORÍA DE LA DEMOCRACIA*, § 49–52 (2006).

² Habermas, *Concluding Comments*, *supra* note 1, at 385.

³ James Bohman, *Survey Article: The Coming of Age of Deliberative Democracy*, 6(4) J. POL. PHIL. 400 (1998); Bächtiger et al., *supra* note 1, at 2; Mathew Ryan & Graham Smith, *Defining Mini-Publics*, in *DELIBERATIVE MINI-PUBLICS. INVOLVING CITIZENS IN THE DEMOCRATIC PROCESS* 9, 9 (Kimmo Grönlund, André Bächtiger, & Maija Setälä eds., 2014).

⁴ Laura Black et al., *Methods for Analyzing and Measuring Group Deliberation*, in *THE SOURCEBOOK FOR POLITICAL COMMUNICATION RESEARCH: METHODS, MEASURES, AND ANALYTICAL TECHNIQUES* 323, 325 (Erik P. Bucy & R. Lance Holbert eds., 2011); Marlène Gerber, *Equal Partners in Dialogue? Participation Equality in a Transnational Deliberative Poll (Europolis)*, 63(1) POL. STUD. 110, 115 (2015).

Scholars measure *popular* deliberation in, roughly, two ways.⁵ First, they contrast results of deliberative exercises against their initial goals, using measures internal to their experiments. This is what Black and her coauthors refer to as *indirect measures*.⁶ As they affirm, indirect measures look for indicators “that deliberation might occur or has occurred,”⁷ and this is usually made either by looking at antecedents (e.g., by determining the extent to which conditions for deliberation are met)⁸ or by focusing on outcomes (e.g., by asking participants to assess their experiences in sophisticated pre- and post-deliberation surveys).⁹ This is a method commonly employed, for instance, in deliberative polls,¹⁰ and national issues forums.¹¹ Both have been applied to different sorts of issues.¹²

By contrast, *direct measures* examine the debate within small groups to determine the “extent to which the discussion corresponds to theoretical conceptions of deliberation.”¹³ These include, *participants assessments*, *integrated case studies*, and *discussion analyses*. Yet, given that the most prominent example of direct measurement is discussion analysis, I will focus on this sort of method.

There are good reasons for focusing on discussion analyses as exemplars of deliberative measurement instruments. They represent, admittedly, a fraction of empirical research on deliberation, as most experiments analyze the influence that participation

⁵ In this, I follow Bächtiger et al., *supra* note 1, at 2, and Black et al., *supra* note 4. However, the taxonomy does not rule out cases where both direct and indirect measures are applied in the same study. An example of such case is Mucciaroni and Quirk’s study of legislation in the US Congress where content analysis and outcome evaluation are intermingled. Gary Mucciaroni & Paul J. Quirk, *Rhetoric and Reality: Going Beyond Discourse Ethics in Assessing Legislative Deliberation*, 4(1) *LEGISPRUDENCE* 35, 39, 43 (2010).

⁶ Black et al., *supra* note 4, at 327, 335–338.

⁷ *Id.* at 335.

⁸ *Id.* at 335–336.

⁹ *Id.* at 327, 336–338.

¹⁰ Lynn M. Sanders, *Poll Envy: An Assessment of Deliberative Polling*, 9(1) *THE GOOD Soc’TY* 9 (1999); James S. Fishkin, *Consulting the Public Through Deliberative Polling*, 22(1) *J. POL’Y ANALYSIS & MGMT.* 128, 128 (2003); BRUCE ACKERMAN & JAMES FISHKIN, *DELIBERATION DAY* § 4 (2004); James Fishkin & Cynthia Farrar, *Deliberative Polling. From Experiment to Community Resource*, in *THE DELIBERATIVE DEMOCRACY HANDBOOK. STRATEGIES FOR EFFECTIVE CIVIC ENGAGEMENT IN THE 21ST CENTURY* 68, 72–73 (John Gastil & Peter Levine eds., 2005); Cynthia Farrar et al., *Disaggregating Deliberation’s Effects: An Experiment Within a Deliberative Poll*, 40 *BRIT. J. POL. S.* 333, 334–336 (2010); David Denver, Gordon Hands, & Bill Jones, *Fishkin and the Deliberative Opinion Poll: Lessons from a Study of the Granada 500 Television Program*, 12(2) *POL. COMM.* 147, 153 (2010); Black et al., *supra* note 4, at 336–337; Simon Niemeyer, *The Emancipatory Effect of Deliberation: Empirical Lessons from Mini-Publics*, 39(1) *POL. & Soc’y* 103, 119 (2011); Pamela Fiber-Ostrow & Sarah A. Hill, *The Deliberative Poll as a Method for Generating Informed Public Opinion*, 12(1) *ANALYSES OF SOC. ISSUES & PUB. POL’y* 151, 152 (2012).

¹¹ John Gastil & James P. Lillard, *The Aims, Methods, and Effects of Deliberative Civic Education Through the National Issues Forums*, 48(3) *COMM. EDUC.* 179, 179–180 (1999); Keith Melville, Taylor L. Willingham, & John R. Dedrick, *National Issues Forums: A Network of Communities Promoting Public Deliberation*, in *THE DELIBERATIVE DEMOCRACY HANDBOOK: STRATEGIES FOR EFFECTIVE CIVIC ENGAGEMENT IN THE 21ST CENTURY* 37, 39 (John Gastil & Peter Levine eds., 2005).

¹² See, e.g., John Uhr, *Parliament and Public Deliberation: Evaluating the Performance of Parliament*, 24(3) *U. NEW S. WALES L.J.* 708 (2001); Denver, Hands, & Jones, *supra* note 10; Farrar et al., *supra* note 10, at 338–339; Niemeyer, *supra* note 10, at 112–118, 119–124.

¹³ Black et al., *supra* note 4, at 326–327.

in institutions like mini-publics exercise on their participants' opinions, using pre- and post-deliberation surveys. Indeed, several experiments set their own deliberative goals and measure the degree to which those goals have been met. Yet, direct measures in general, and discussion analyses in particular, represent the mainstream approach in the literature in terms of comprehensive measurement.¹⁴

Different, more or less updated versions of the coding schemes I analyze below are used in recent empirical studies. Consider, for example, the recent application of Steenberger et al.'s Discourse Quality Index (DQI) to deliberative polling in order to determine levels of deliberation among the different participants in the Europolis project.¹⁵ Ryan and Smith approve of their use when they say that "there is a promise on the further application of the Discourse Quality Index (DQI) on the deliberations of participants in . . . Europe-wide Deliberative Polls."¹⁶

Pedrinì also developed a coding scheme that "attempts to contribute to the re-development of the DQI, which—in its original conception—has not comprised expanded notions of deliberative quality such as storytelling."¹⁷

Additionally, Jaramillo has elaborated a coding scheme based on and aiming to improve the DQI. In her words, "when it comes to quick give-and-takes in small group discussions, the DQI has its shortcomings, which we attempt to remedy with the concept of Deliberative Transformative Moments (DTM). This concept shall not replace but amend the DQI."¹⁸ The use of this amended version of the DQI is not limited to this work. As it happens, the complete elaboration of this content analysis tool is made in a recent publication.¹⁹

Likewise, the DQI has influenced the V-Dem project, a state-of-the-art research project on democratic quality, broadly understood. The project measures the deliberative quality of different polities by examining discourses according to a series of categories clearly inspired by DQI—like coding schemes, namely, reasoned justification, common good, respect for counterarguments, range of consultation among group members, level of social engagement, the particularistic or public nature of social goods, and whether welfare programs are means-tested or universalistic.²⁰

¹⁴ For examples of an alternative approach, see John Gastil, Katie Knobloch, & Meghan Kelly, *Evaluating Deliberative Public Events and Projects*, in *DEMOCRACY IN MOTION: EVALUATING THE PRACTICE AND IMPACT OF DELIBERATIVE CIVIC ENGAGEMENT* 205, 225–229 (Tina Nabatchi et al. eds., 2012); Harmut Wessler, *Investigating Deliberativeness Comparatively*, 25(1) *POL. COMM.* 1 (2008).

¹⁵ Pierangelo Isernia et al., *Toward a European Public Sphere: The Europolis Project*, in *IS EUROPE LISTENING TO US? SUCCESSES AND FAILURES OF EU CITIZENS CONSULTATIONS* 79 (Raphaël Kies & Patrizia Nanz eds., 2012).

¹⁶ Ryan & Smith, *supra* note 3, at 23.

¹⁷ Seraina Pedrinì, *Deliberative Capacity in the Political and Civic Sphere*, 20(2) *SWISS POL. SCI. REV.* 263, 264 (2014).

¹⁸ María C. Jaramillo, *Deliberative Transformative Moments: A New Concept as Amendment to the Discourse Quality Index*, 10(2) *J. PUB. DELIBERATION* 1, 1 (2014).

¹⁹ JÜRGEN STEINER ET AL., *DELIBERATIVE POLITICS IN ACTION. ANALYZING PARLIAMENTARY DISCOURSE* (2004).

²⁰ Michael Coppedge et al., *Vdem Varieties of Democracy Project—Vdem Codebook v.6*, in *V-DEM: VARIETIES OF DEMOCRACIES* 192–197 (Mar. 6, 2016), https://www.v-dem.net/media/filer_public/d1/24/d124efd5-7ff5-4175-a1ed-f294984084d0/v-dem_codebook_v6.pdf.

Moreover, Gerber, Bächtiger and others have recently employed modified versions of the DQI to explore “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 or arguments.”²¹

Discussion analyses scrutinize communication processes in deliberative sessions, typically through the examination of records, transcripts, and videos, or in the verbatim generated record of online discussions.²² There are two subcategories of discussion analysis that can be used complementary in a single experiment: *micro-* and *macro-analytic* approaches.²³ *Micro* approaches evaluate the content of discussions, taking texts as elements of analysis, from which researchers try to make valid and meaningful inferences. Such examination can be quantitative, by calculating the number of words or expressions uttered by participants, or by measuring the amount of time each participant has spoken,²⁴ but it can also be qualitative, content-based, where meaning springs from interpretative procedures analysts apply to the discussions.

Content analysis is typically made by first developing a codebook containing the elements to be analyzed during the deliberation. Two or more coders are then trained to analyze the discussion and to apply the scheme. It is important that the analysts agree on the codes they give to each assessed unit, for the more they agree the more reliable the coding scheme and the measurement is expected to be. In turn, in *macro* approaches, coders are asked to make “summary judgments of the discussions as a whole.”²⁵

There are several instruments discussed in the literature so, although there are others available,²⁶ I discuss the following four, as they are representative of the ways in which content analysis work: Steenbergen et al.’s DQI,²⁷ Bächtiger et al.’s DQI 2.0,²⁸ Black et al.’s measurement of online discussions,²⁹ and Stromer-Galley’s coding scheme.³⁰

²¹ Marlene Gerber et al., *Deliberative Abilities and Influence in a Transnational Deliberative Poll (EuroPolis)*, BRIT. J. POL. SCI. (Sept. 15, 2016), available at <https://doi.org/10.1017/S0007123416000144>.

²² Black et al., *supra* note 4, at 326.

²³ Bächtiger et al., *supra* note 1, at 2; Black et al., *supra* note 4, at 327–334.

²⁴ Bächtiger et al., *supra* note 1; Gerber, *supra* note 4, at 114.

²⁵ Bächtiger et al., *supra* note 1, at 2.

²⁶ E.g., Jürgen Gerhards, *Diskursive vs liberale Öffentlichkeit: Eine empirische Auseinandersetzung mit Jürgen Habermas*, 49(1) KÖLNER ZEITSCHRIFT FÜR SOZIOLOGIE UND SOZIALPSYCHOLOGIE 1 (1997); Katharina Holzinger, *Verhandeln statt Argumentieren oder Verhandeln durch Argumentieren? Eine empirische Analyse auf der Basis der Sprechaktheorie*, 42(3) POLITISCHE VIERTELJAHRESSCHRIFT 414 (2001); Davi Janssen & Raphaël Kies, *Online Forums and Deliberative Democracy*, 40 ACTA POL. 317 (2005); RAPHÉL KIES, *PROMISES AND LIMITS OF WEB-DELIBERATION* (2010).

²⁷ Marco R. Steenbergen et al., *Measuring Political Deliberation: A Discourse Quality Index*, 1 COMP. EUR. POL. 21 (2003).

²⁸ Bächtiger et al., *supra* note 1.

²⁹ Laura Black et al., *Self-Governance Through Group Discussion in Wikipedia: Measuring Deliberation in Online Groups*, 42(5) SMALL GROUP RES. 595 (2011).

³⁰ Jennifer Stromer-Galley, *Measuring Deliberation’s Content: A Coding Scheme*, 3(1) J. PUB. DELIBERATION 1 (2007).

Steenbergen et al.'s DQI is the most discussed available instrument.³¹ It consists of a coding scheme based on the following normative tenets of Habermasian discourse ethics: open participation, justification, search for the common good, respect, constructive politics, and authenticity.³²

First, "there should be open participation," so that everyone is allowed to introduce assertions into the debate and express their attitudes, desires, and needs without coercion.³³

Second, deliberation demands that claims, assertions, desires, needs, and so on, be justified. Those justifications should be logically coherent, coherence measured by the degree to which conclusions follow from premises; "the tighter the connection between premises and conclusions, the more coherent the justification is and the more useful for deliberation."³⁴

Third, participants "should consider the *common good*,"³⁵ by which they mean that "there should be a sense of empathy, other-directness, or solidarity that allows the participants to consider the well-being of others and of the community at large."³⁶ Self-interest, though not ruled out, is accepted if it is demonstrated that "it is compatible with or contributes to the common good." Provided that the notion of common good can be understood in different ways, they think its relation with self-interest can "be expressed through [Rawls'] difference principle, that is, accepted to the extent it serves the least advantaged."³⁷

Fourth, there should be respect among participants as it is, according to them, a prerequisite for serious listening, something essential for deliberation. In turn, respect can be expressed toward groups, demands, and counterarguments, where the latter is, in particular, "a necessary condition for weighing of alternatives."³⁸

Fifth, deliberation should aim at *constructive politics*, that is, Steenbergen et al.'s understanding of consensus. Consensus, they note, is not an absolute necessity, as it is often not even possible in politics. Yet, for them it is important "that participants in a discourse . . . at least attempt to reach mutually acceptable compromise solutions, since this is the only way in which universalism can be attained."³⁹

Finally, their scheme considers *authenticity*, "which is the absence of deception in expressing intentions."⁴⁰ They concede this is an important element of deliberation. Yet based on the difficulty of judging when exactly a certain speech is authentic or deceptive, they consider their measurement instrument would be bound to "introduce

³¹ Gold, Holzinger, & Rohrdantz, *supra* note 1, at. 1–2. Here I comment on Steenbergen et al., *supra* note 27. The full elaboration of the DQI is in Steiner et al., *supra* note 19, at 43–73.

³² Steenbergen et al., *supra* note 27, at 25.

³³ *Id.*

³⁴ *Id.*

³⁵ *Id.* (emphasis in the original).

³⁶ *Id.* at 26.

³⁷ *Id.*

³⁸ *Id.*

³⁹ *Id.*

⁴⁰ *Id.*

large amounts of (possibly systematic) measurement error.” For this reason, they do not include authenticity in their coding scheme.⁴¹

These normative tenets ground the coding scheme, which is divided into three parts: the unit of analysis, the coding categories, and the index itself.

The unit of analysis is speech, understood as a public discourse made by an individual delivered at a particular point in a debate. The speech is then broken down into smaller units whose parts are considered as relevant if they contain demands, or irrelevant if who utters them makes no demand—as, for example, when someone makes clarifications or asks for them.⁴²

As for the coding categories, they are applied to speeches containing relevant parts, and measure participation, the levels of justification given for arguments, respect for others’ demands, and constructive politics.⁴³ The procedure for using these codes in concrete examples of deliberation requires two or more coders. Roughly, each of them must first read through the entire debate individually and code the relevant speeches by assigning them numerical value. Then, they come together to compare codes. If coders disagree, they read through the speech again and discuss the merits of the codes they assigned to those speeches. In the end, coders must deliberate in order to convince each other of the correctness of their evaluation.⁴⁴

Bächtiger et al. designed a redeveloped version in order to cope with the shortcomings of the original DQI.⁴⁵ They first proceed by defining two groups of normative standards they call type I, “which captures rational discourse” and type II deliberation, “which measures alternative forms of communication such as ‘story-telling’ and ‘deliberative negotiations.’”⁴⁶ Type I deliberation “embodies the idea of rational discourse, focuses on deliberative intent and the related distinction between communicative and strategic action, and has a strong procedural component.”⁴⁷ Type II deliberation “generally involves more flexible forms of discourse, more emphasis on outcomes versus process, and more attention to overcoming ‘real-world’ constraints on realizing normative ideals.”⁴⁸

Type I standards measure equality, justification rationality, common good orientation, respect and agreement, interactivity, constructive politics, and sincerity. I will not

⁴¹ *Id.* Likewise, Bächtiger et al. drop the criterion of sincerity or truthfulness from their deliberation measuring standards, given that “true preferences are not directly observable.” Bächtiger et al., *supra* note 1, at 6

⁴² Steenbergen et al., *supra* note 27, at 27.

⁴³ *Id.* at 27–30; Martin King, *A Critical Assessment of Steenbergen’s Discourse Quality Index*, 1 *ROUNDHOUSE: J. CRITICAL THEORY & PRAC.* 1, 2 (2009).

⁴⁴ Steenbergen et al., *supra* note 27, at 31.

⁴⁵ Bächtiger et al., *supra* note 1, at 2.

⁴⁶ *Id.* at 4; Bächtiger et al., *Disentangling Diversity in Deliberative Democracy: Competing Theories, Their Blind Spots and Complementarities*, 19(1) *J. POL. PHIL.* 32, 33 (2010).

⁴⁷ *Id.*

⁴⁸ *Id.* See also Michael Morrel, *Participant Bias and Success in Deliberative Mini-Publics*, in *DELIBERATIVE MINI-PUBLICS INVOLVING CITIZENS IN THE DEMOCRATIC PROCESS* 151, 159 (Kimmo Grönlund, André Bächtiger, & Maija Setälä eds., 2014).

comment on justification rationality and common good orientation, as they remain the same as in the original DQI.⁴⁹

On the other hand, type II standards include less rational forms of discourse.⁵⁰ Following Mansbridge,⁵¹ the new version of the DQI includes two elements: storytelling and bargaining. Storytelling is measured by determining whether participants use personal narratives or experiences.⁵² As to bargaining or “deliberative negotiations,” theoretically, these authors follow Mansbridge, and distinguish between deliberative and non-deliberative negotiations.⁵³ Empirically, they follow Holzinger,⁵⁴ and try to capture different forms of bargaining, counting whether a speech contains threats or promises.⁵⁵

Like the rest of the measuring instruments, the authors of this version of the DQI create a codebook and apply these standards to actual conversations, assigning codes for each unit of analysis.⁵⁶

Online deliberation has prompted the emergence of measuring instruments in online groups. Black et al. investigated, for example, policymaking discussions on Wikipedia “with an eye toward informing future research on virtual teams and online discussion.”⁵⁷ They developed a content analysis scheme to measure “the extent to which policy-making discussions adhere to idealized models of high-quality group deliberation.”⁵⁸

Following Gastil and Black,⁵⁹ Black et al. single out five analytic aspects of deliberation: creating an information base, prioritizing key values, identifying solutions, weighing those solutions, and making best decisions. Deliberation also involves four social components, namely, equality of speaking opportunities, mutual comprehension, consideration of others, and respect. These analytical and social dimensions “form the foundation for the coding scheme used in this research.”⁶⁰

The unit of analysis was the discussion post, examined in the order they appeared on the talk page. In addition to individual posts, “discussion threads in their entirety were coded for overall summary judgments.”⁶¹ Researchers were in charge of coding the posts on eight of the nine aforementioned dimensions. Equality was not included

⁴⁹ Bächtiger et al., *supra* note 1, at 5–6.

⁵⁰ *Id.* at 7.

⁵¹ Jane Mansbridge, “*Deliberative Democracy*” or “*Democracy Deliberation*,” in *DELIBERATION, PARTICIPATION AND DEMOCRACY: CAN THE PEOPLE GOVERN?* 251 (Shawn W. Rosenberg ed., 2007).

⁵² Bächtiger et al., *supra* note 1, at 8.

⁵³ *Id.*

⁵⁴ See Holzinger, *supra* note 26.

⁵⁵ Bächtiger et al., *supra* note 1, at 8.

⁵⁶ For an example of the application of this scheme, see André Bächtiger et al.’s project, “Potential for Deliberation Among EU Citizens” (2010–2013).

⁵⁷ Black et al., *supra* note 29, at 596.

⁵⁸ *Id.*

⁵⁹ John Gastil & Laura Black, *Public Deliberation as the Organizing Principle of Political Communication Research*, 4(1) J. PUB. DELIBERATION 1, 1 (2008).

⁶⁰ Black et al., *supra* note 29, at 597–599.

⁶¹ *Id.* at 607.

in coding, but “it was captured in global ratings that coders made for each discussion thread.”⁶²

Stromer-Galley also designed a content analysis scheme to measure the quality of political deliberation in face-to-face and online groups. She understands deliberation as “a process whereby groups of people, often ordinary citizens, engage in reasoned opinion expression on a social or political issue in an attempt to identify solutions to a common problem and to evaluate those solutions.”⁶³ Her coding scheme considers six elements: “reasoned opinion expression, references to external sources when articulating opinions, expressions of disagreement and hence exposure to diverse perspectives, equal levels of participation during the deliberation, coherence with regard to the structure and topic of deliberation, and engagement among participants with each other.”⁶⁴

Stromer-Galley developed her coding scheme from these theoretical elements of deliberation. For this, she developed a content analysis codebook whose units for coding “were identified at the level of the *thought*,”⁶⁵ defined as “an utterance (from a single sentence to multiple sentences) that expresses an idea on a topic.”⁶⁶

2.1. Limitations

I now criticize these measures. Before I detail more general limitations shared by all the coding schemes described above, some words about the specific shortcomings evinced by the DQI are in order, given its utility and widespread use.⁶⁷

I will not delve too much into shortcomings that are already in the literature. Here, it is enough to remember that Steenbergen et al. themselves have noted problems derived from their decision to drop authenticity from the scheme,⁶⁸ from the narrowness of its Habermasian theoretical foundations,⁶⁹ from the fact that the scheme does not measure non-textual communication, and from its inability to capture participation in a more substantive way than by coding whether participants are interrupted.⁷⁰ Additionally they have acknowledged that the scheme’s theoretical foundations leave little room for other conceptual developments in deliberative theory.⁷¹

Additionally, the DQI has been criticized for its lack of cut or threshold values determining when a given normative goal is met, which limits the index to a mere

⁶² *Id.* at 608.

⁶³ Stromer-Galley, *supra* note 30, at 4.

⁶⁴ *Id.*

⁶⁵ *Id.* at 9 (emphasis in the original).

⁶⁶ *Id.* at 8.

⁶⁷ For examples of its use, see Steenbergen et al., *supra* note 27, at 30–41; Steiner et al., *supra* note 19; Michael Coppedge et al., *Conceptualizing and Measuring Democracy: A New Approach*, 9(2) PERSP. ON POL. 247 (2011); JÜRGEN STEINER, *THE FOUNDATIONS OF DELIBERATIVE DEMOCRACY: EMPIRICAL RESEARCH AND NORMATIVE IMPLICATIONS* (2012); Isernia, et al., *supra* note 15.

⁶⁸ Steenbergen et al., *supra* note 27, at 43.

⁶⁹ *Id.*

⁷⁰ *Id.* at 43–44. Likewise, King, *supra* note 43, at 4–5.

⁷¹ Bächtiger, et al., *supra* note 1, at 3.

comparative function,⁷² as well as for the fact that it only assesses the quality of entire debates, when it is hardly the case that deliberation is present throughout the entire discussion.⁷³

Furthermore, King notes that the DQI allows incoherent comparisons and balancing between different aspects of speech: “For example, to allow statements declaring that ‘the justifications were inferior, but the participation was good, so overall the discourse was’, does not seem a coherent or desirable usage of discourse ethics theory.”⁷⁴

Finally, others have pointed to practical problems in using the DQI like, for example, the amount of time needed to undertake these experiments.⁷⁵

But the DQI is not the only coding scheme facing problems. Some scholars have also highlighted several limitations affecting all the preceding instruments of deliberative measurement.⁷⁶ Yet, the following criticisms have not been fully elaborated.

First, at the level of ideal theory, there is, generally, a sort of fixation with the Habermasian version of deliberative communication that is not adequately justified. For example, Steenbergen et al. argue that they selected a theory “most closely associated with Habermas’ . . . discourse ethics.”⁷⁷ Stromer-Galley adopts a definition of deliberation that “aligns most closely with that of Schudson,⁷⁸ Habermas,⁷⁹ as well as Gastil.”^{80,81} Also, Gold et al.’s coding scheme starts “from the ideas presented by Habermas in his work on communicative action and discourse ethics. That is,” they claim, “conceptually, we rely on a Habermasian approach to deliberation.”⁸² They thus define deliberation as “a communicative process that aims at taking a decision (or recommendation) on collectively binding rules or public projects. The substantive goal is to achieve the common good and universality of rules.”⁸³

This approach needs further justification. Deliberative democracy is a theoretical construction to which Habermas contributed, but his is not the only understanding of what deliberative democracy is, and the choice for his over others’ versions of the theory begs for explanations.

Hence, we are in the position to ask: why only Habermas and not any other deliberative democrat(s) whose contribution had been significant for the development of the theory, contribution(s) that could reasonably lead to different or complementary measuring standards. Kies rightly observes that empirical approaches based on, for

⁷² In the same vein, King, *supra* note 43, at 5.

⁷³ Bächtiger et al., *supra* note 1, at 3.

⁷⁴ King, *supra* note 43, at 4.

⁷⁵ Gold et al., *supra* note 1, at 2.

⁷⁶ E.g., Ricardo Mendonça, *Assessing Some Measures of Online Deliberation*, 9(3) BRAZILIAN POL. SCI. REV. 88, 97–105 (2015).

⁷⁷ Steenbergen, et al., *supra* note 27, at 23, 26.

⁷⁸ Michael Schudson, *Why Conversation Is Not the Soul of Democracy*, 14 CRITICAL STUD. IN MASS COMM. 297 (1997).

⁷⁹ JÜRGEN HABERMAS, *THEORY OF COMMUNICATIVE ACTION: REASON AND THE RATIONALIZATION OF SOCIETY* (1981).

⁸⁰ JOHN GASTIL, *BY POPULAR DEMAND* (2000).

⁸¹ Stromer-Galley, *supra* note 30, at 4.

⁸² Gold et al., *supra* note 1, at 2.

⁸³ *Id.*

example, Young's promotion of social justice are likely to insist on the criterion of inclusion. Those inspired by Gutmann and Thompson's notion of disagreement will put the accent in the criterion of respect. Those "inspired by the theories of Sennet or Sunstein [are] more likely to test whether online public spaces have allowed for confrontation of a multiplicity of unexpected and spontaneous opinions."⁸⁴ Simply claiming that internal consistency is the reason to choose one and not others⁸⁵ bypasses important contributions to the evolution and a full understanding of deliberative democratic theory.

An additional problem is raised by this fixation with Habermas: these authors use expressions like "communicative action," "discursive rationality," "discourse ethics," "deliberative democracy," etc., haphazardly, without providing specific paragraphs and/or page numbers, failing to show that these notions assume different meanings and are applied to different concerns in different periods of Habermas's intellectual development.

This is particularly stark in the use of the concepts of discourse ethics and discourse theory. The first is a concept that Habermas used during early stages of his work, where communicative rationality was meant to be employed by individuals who engaged in dialogue, where claims were addressed by each other assuming substantive normative presuppositions absent of instrumental rationality. Discourse ethics was independent of the political and legal realms, as "[i]t is by no means self-evident that rules which are unavoidable *within* discourses can also claim to be valid for regulating action *outside* of discourses."⁸⁶

Habermas's explicit concern with this application of his theory to the political realm came later with his construction of a *theory of discourse*, a rather different understanding and employment of discursive rationality where the ethical concern was expanded to other fields of practical reason, and where the justification of norms includes not only ethics but law and politics as well. We find evidence of this change in Habermas's extension of principle (D), which in *Moral Consciousness and Communicative Action* "contained the distinctive idea of an *ethics of discourse*,"⁸⁷ whereas, for example, in *Between Fact and Norms*, "it lies at a different level than the distinction between moral and ethical discourse. As a principle for the impartial justification of norms in general, (D) also underlies both morality and law."⁸⁸ In fact, Habermas has been criticized on the grounds that in making this move toward a theory of discourse, he neglected or diluted the ethical dimension of communicative rationality.⁸⁹ A proper application of Habermas's theory ought to have these differences under consideration.

⁸⁴ Kies, *supra* note 26, at 41 (references omitted).

⁸⁵ Steenbergen et al., *supra* note 27, at 23.

⁸⁶ JÜRGEN HABERMAS, *MORAL CONSCIOUSNESS AND COMMUNICATIVE ACTION* §§ 85–86 (1990) (emphasis in the original).

⁸⁷ *Id.* at 66 (emphasis added).

⁸⁸ William Rehg, *Translator's Introduction*, in *BETWEEN FACTS AND NORMS: CONTRIBUTIONS TO A DISCOURSE THEORY OF LAW AND DEMOCRACY* §§ ix, xxvi (1996).

⁸⁹ Adela Cortina, *La política deliberativa de Jürgen Habermas: Virtualidades y límites*, 144 *REVISTA DE ESTUDIOS POLÍTICOS* 169, 172 (2009).

This shift of perspective has theoretical consequences, most prominently, for the role played by law in the configuration of a deliberative political system. During Habermas's early period, law was part of the system, and juridification a pathological manifestation of legality, bureaucratizing, and invading myriad forms of social life, turning citizens into clients of bureaucracies.⁹⁰ Things are different in Habermas's later works, which he developed from the late 1980s onward,⁹¹ where law is conceived of as an articulator of the subsystems that together form society, a translator of their different discourses. The bureaucratic and economic system, as well as civil society are now related in ways that admit both communicative and instrumental rationality, as well as forms of reasoning that are now accepted to the extent they are translated into decisions through a legitimate legal system that is the result of a deliberative procedure.⁹²

These differences show that authors may be promising one thing, that is, to measure political communication or political deliberation, and doing another in an important way, namely, measuring the communicative rationality of different forms of discourse from the point of view of discourse ethics, or vice versa. One could object to this criticism on the basis that the original endeavor—measuring the quality of deliberation—remains firm regardless of this conceptual distinction. Just replace, our objector could say, discourse ethics with theory of discourse, or some other expression which do justice to the evolution of Habermas's approach. After all, Habermas himself has expressed satisfaction regarding the way in which his theory has been operationalized by social scientists.⁹³

Nonetheless, it still remains to be shown that the wording employed in the formulation of the theoretical grounds of their experiments does not hide deeper theoretical misconceptions or misused terminologies. My take on this matter is that deliberative democrats may be using terms without reflecting properly about concepts.⁹⁴

This is not, however, the only limitation these schemes face. Another concern is the difficulty of measuring authenticity. I have mentioned that coding schemes generally drop this category from the analysis, for it is an element that may only be manifested in *foro interno*, limiting social scientists' abilities to codify utterances as sincere or insincere.⁹⁵ Kies, for example, after acknowledging these difficulties, asks whether this means that we should "ignore the criterion as most of the empirical research does?"⁹⁶ He answers in the negative affirming that "the criterion of sincerity with the one of empathy together form a cardinal evaluative criterion of deliberation." He thus concludes that "[t]his means that if sincerity is absent a debate cannot be considered as deliberative even if all the other deliberative criteria score high."⁹⁷

⁹⁰ Habermas, *supra* note 79, § 3.4.

⁹¹ LASSE THOMASSEN, HABERMAS: A GUIDE FOR THE PERPLEXED § 85 (2010).

⁹² Habermas, *Concluding Comments*, *supra* note 1, at 165.

⁹³ *Id.* at 384. Specifically, by Janssen & Kies, *supra* note 26.

⁹⁴ Johan Olsthoorn, *Conceptual Analysis*, in *METHODS IN ANALYTICAL POLITICAL THEORY* 153 (Adrian Blau ed., 2017).

⁹⁵ Steenbergen et al., *supra* note 27, at 26; King, *supra* note 43, at 3; Bächtiger et al., *supra* note 1, at 6.

⁹⁶ Kies, *supra* note 26, at 52.

⁹⁷ *Id.* Likewise, Lincoln Dahlberg, *The Internet of Democratic Discourse: Exploring the Prospects of Online Deliberative Forums Extending the Public Sphere*, 4(4) INFO., COMM. & SOC'Y 615, 626 (2004).

I disagree with the last part of Kies's argument. As it stands, his conclusion is too strong. Even though we lack reliable instruments for the measurement of sincerity, deliberative democrats have been adamant that deliberators face constraints intrinsic to the deliberative process that compensate, as it were, for shortages in truthfulness. Deliberative settings "can shape outcomes independently of the motives of the participants."⁹⁸ This, what Elster coins the "civilizing force of hypocrisy," thus imposes an *imperfection constraint* that avoids a perfect coincidence between private interest and impartial argument.⁹⁹ But it also imposes a *consistency constraint*: "[o]nce a speaker has adopted an impartial argument because it corresponds to his interest or prejudice, he will be seen as opportunistic if he deviates from it when it ceases to serve his needs."¹⁰⁰ In spite of the fact that we should pursue sincerity for reasons related to transparency,¹⁰¹ the notion of the civilizing force of hypocrisy argues against Kies's contention that low scores on sincerity somehow cancel or trump high scores on other criteria, because the procedure itself gives participants incentives to wield neutral rather than selfish arguments if they want to achieve their self-interested goals. Thus, procedures may still be deliberative when participants are being insincere.

Another problem is that the coding schemes measure equality either by counting the number of words or by determining whether participants have been interrupted. This is problematic, first, because the meaning of equality in deliberative democracy is a contested issue, more complex than what counting words or interruptions may account for, for the literature on equality in deliberative democracy includes formal and material aspects that are not covered by the coding schemes.¹⁰² Second, because there is no direct correlation between equality and the number of words spoken. It could be the case that individuals who speak a small number of words may have a dominant position in the group and only need to utter few expressions to impose their preferences. Conversely, people can utter many words without saying anything meaningful.

Additional limitations arise from the application of these coding schemes to deliberative systems.¹⁰³ On the one hand, it remains unexplored how to measure the quality

⁹⁸ Jon Elster, *Deliberation and Constitution Making*, in *DELIBERATIVE DEMOCRACY* 97, 103 (Jon Elster ed., 1998).

⁹⁹ *Id.*

¹⁰⁰ *Id.*

¹⁰¹ AMY GUTMANN & DENNIS THOMPSON, *DEMOCRACY AND DISAGREEMENT* §§ 100–101 (1996); Martí, *supra* note 1, at 93; Immanuel Kant, *Toward Perpetual Peace*, in *TOWARD PERPETUAL PEACE AND OTHER WRITINGS ON POLITICS, PEACE, AND HISTORY* 67 (Pauline Kleingeld ed., 2006).

¹⁰² Martí, *supra* note 1, at 95–96.

¹⁰³ Jane Mansbridge, *Everyday Talk in the Deliberative System*, in *DELIBERATIVE POLITICS: ESSAYS ON DEMOCRACY AND DISAGREEMENT* 211 (Stephen Macedo ed., 1999); John Dryzek, *Global Democratization: Soup, Society, or System?*, 25(2) *ETHICS & INT'L AFF.* 211 (2011); JOHN PARKINSON, *DEMOCRATIZING DELIBERATIVE SYSTEMS* (2012); James Bohman, *Representation in the Deliberative System*, in *DELIBERATIVE SYSTEMS: DELIBERATIVE DEMOCRACY AT THE LARGE SCALE. THEORIES OF INSTITUTIONAL DESIGN* 72 (John Parkinson & Jane Mansbridge eds., 2012); Jane Mansbridge et al., *A Systemic Approach to Deliberative Democracy*, in *DELIBERATIVE SYSTEMS: DELIBERATIVE DEMOCRACY AT THE LARGE SCALE. THEORIES OF INSTITUTIONAL DESIGN* 1 (John Parkinson & Jane Mansbridge eds., 2012); Jonathan Kuyper, *Democratic Deliberation in the Modern World: The Systemic Turn*, 27(1) *CRITICAL REV.* 49 (2015).

of the parts, the subsystems, and the system of deliberation itself.¹⁰⁴ On the other, I do not see how, in the face of this lack of mechanisms, one could apply the coding schemes I discussed above from a systemic perspective. More research is needed in order to measure the quality and quantity of deliberative systems using new instruments or adapting the available ones.

These considerations show that measuring the quality and quantity of deliberation in real-world fora is a pressing, however, limited, project. There is a disconnection between political theory and social science that leaves a gap in terms of the theoretical foundations of the measuring schemes. This could be the manifestation of an old problem: either our normative conceptions of deliberative democracy are too ideal to be turned into empirical models, or our practices fail to live up to our principles. It could be a bit of both, but, given the limitations here explored, the burden is more on social science than on political theory to work out ways in which deliberation can be measured more appropriately.

One key message of this section is that our evaluations of the quality and quantity of deliberative sites are to be accepted with a grain or two of salt. In spite of the utility and widespread use of the analyzed measuring schemes, there is still work to be done in order to say, with full confidence, that the normative desiderata underpinning deliberative democratic theory are realized (or not) in these or those sites, under these or those conditions. These schemes are the best we have, but they are far from being perfect. The main conclusion I can draw for my purposes here is that idealizations of deliberative sites are to be avoided.

3. Judicial deliberation

And yet, there is one institution that is portrayed by legal and political scholars in commendatory terms regarding their deliberative virtues, namely, the judiciary. In spite of the conclusions to which I arrived in the previous section and despite the fact that, “[e]mpirical evidence about the behavior of judges is almost never offered to support the contrast [between *popular* and *judicial* institutions],”¹⁰⁵ courts are held in high regard in terms of their deliberative capacities.

The remainder of this article thus focuses on judicial deliberation. It argues that a combined reading of the findings made in the previous section and the ones made in this part shows that idealizations of the judiciary in terms of its deliberative capacity are to be avoided.

As it happens, courts have been objects of such idealization, despite two facts: first, as I have showed in the previous section, our capacity to measure the quality with which agents deliberate is still limited. Second, judicial proceedings and rules do not enable democratic deliberation in the way demanded by deliberative democratic theory. This section elaborates on this second problem.

¹⁰⁴ Mendonça, *supra* note 76, at 101.

¹⁰⁵ Gutmann & Thompson, *supra* note 101, at 45.

It is indeed perplexing that a number of prominent political and legal scholars have expressed commendatory opinions for the judiciary, and for some features of the courtroom that would allegedly turn it into an ideal place for deliberation.¹⁰⁶ Consider, for example, Gutmann and Thompson's statement:

[m]any constitutional democrats focus on the importance of extensive moral deliberation within one of our democratic institutions—the Supreme Court. They argue that judges cannot interpret constitutional principles without engaging in deliberation, not least for the purpose of constructing a coherent view out of many moral values that our constitutional tradition expresses.¹⁰⁷

Or, consider Sen's opinion:

images of honest deliberation and of intellectual exchange have prompted some political philosophers and legal theorists to cite the judiciary as an example for other branches to follow. . . . They further have assumed the judiciary is best equipped to realize deliberative ideals, and that it can provide an example for the less deliberative branches to follow.¹⁰⁸

Moreover, Ferejohn and Pasquino avow that some scholars claim that if “we are trying to locate the institutions where reasoning and deliberation play an important role in public life, it is apt to begin with courts and especially with courts dealing with constitutional issues.”¹⁰⁹

Think of Rawls's portrayal of the US Supreme Court as the exemplar of public reason and its discourse as the best instance of the language deliberators should use given the expected neutrality of public reason, “carefully eschewing reference to citizens' diverse comprehensive worldviews, while nevertheless rendering decisions based on fundamental political values shared by all reasonable citizens.”¹¹⁰

Additionally, according to Eisgruber, the court ought to be understood as an institution “well-shaped to speak on behalf of the people about questions of moral and political principle,” to “represent the people's convictions about what is right.”¹¹¹ He argues

¹⁰⁶ Maya Sen, *Courting Deliberation: An Essay on Deliberative Democracy in the American Judicial System*, 27 NOTRE DAME J.L. ETHICS & PUB. POL'Y 303, 304 (2013); CHRISTOPHER ZURN, DELIBERATIVE DEMOCRACY AND THE INSTITUTIONS OF JUDICIAL REVIEW §§ xiii–xiv, 167, 172 (2007); Donald Bello Hutt, *Deliberation and Courts: The Role of the Judiciary in a Deliberative System*, 64(3) THEORIA: J. SOC. & POL. THEORY 77 (2017).

¹⁰⁷ Gutmann & Thompson, *supra* note 101, at 45.

¹⁰⁸ Sen, *supra* note 106, at 304.

¹⁰⁹ John Ferejohn & Pasquale Pasquino, *Constitutional Courts as Deliberative Institutions: Toward and Institutional Theory of Constitutional Justice*, in CONSTITUTIONAL JUSTICE: EAST AND WEST 21, 22 (Wojciech Sadurski ed., 2002).

¹¹⁰ ZURN, *supra* note 106, at 167; JOHN RAWLS, POLITICAL LIBERALISM §§ 231–240 (1996). Portraying Rawls as a deliberative democrat is reasonable, although admittedly contestable. For accounts in favor of his inclusion, see Samuel Freeman, *Deliberative Democracy: A Sympathetic Comment*, 29(4) PHIL. & PUB. AFF. 371, 382 (2000); Joshua Cohen, *For a Democratic Society*, in THE CAMBRIDGE COMPANION TO RAWLS 86 (Samuel Freeman ed., 2003); John Dryzek, *Rhetoric in Democracy: A Systemic Appreciation*, 38(3) POL. THEORY 319, 325 (2010); Morrel, *supra* note 48, at 159. Moreover, and while critical of them, Cohen's influential conception of deliberative democracy, clearly springs from Rawlsian roots. See Joshua Cohen, *Deliberation and Democratic Legitimacy*, in DELIBERATIVE DEMOCRACY 67 (James Bohman & William Rehg eds., 1997).

¹¹¹ CHRISTOPHER EISGRUBER, CONSTITUTIONAL SELF-GOVERNMENT §§ 3, 5 (2001).

that “[j]udges are supposed to respond to reasons, not preferences. The structure of federal judicial institutions, and specially of the Supreme Court, makes it likely that judges will be disinterested and hence capable of acting on the basis of reasons rather than interests.”¹¹²

Also, in *Between Facts and Norms*, Habermas avers that constitutional courts are instrumental for meeting and securing the constitutional conditions of deliberative politics. In his view, courts act as impartial referees between citizens and their representatives, ensuring that public opinion is directed into the legally structured public sphere by keeping open the channels of political change, guaranteeing respect for the individuals’ legal, social, and political rights.¹¹³

As a final example, consider Alexy’s notion of discursive representation. Courts, in Alexy’s view, help justify democratic government by representing citizens, not through the traditional political notion of representation, but through the rationality of the procedures by which correct rules and principles are identified and deliberatively applied to individuals.¹¹⁴

The findings discussed in Section 2 show that there are several untested and untestable assumptions in these commendatory opinions. They warn us against being too enthusiastic regarding the deliberative qualities of any fora; our capacity to test those qualities is simply too far from perfect to speak about ideal deliberative sites, courts included. The coding schemes reviewed in the preceding section have been used to evaluate a vast and diverse array of deliberative settings. Yet, there is a staggering shortage in the number of experiments measuring how and to what extent deliberation unfolds inside the courtroom. Inasmuch as courts are affected by our limited capacity to measure how and how much deliberation takes place, and that no line of research has systemically analyzed how and how much courts deliberate, it becomes surprising that courts are held in such a high regard.

Hence, the question raised by the opinions described is: how do we know that it is in fact the case that courts are, not only deliberative institutions but *ideal* deliberative institutions? The limitations of measuring instruments explored above tell us not to idealize *any* fora, the literature on the matter is scarce, the measuring schemes have not been applied to courtrooms, and there are no studies or analogous mechanisms telling us how judges deliberate in practice. In view of these shortages, the intuition that judges deliberate before they decide is difficult to justify. Any evidence we may find, e.g., judges’ opinions or interviews, is likely to be contingent.

But there is one more reason why we ought to be wary, and indeed avoid portraying the judiciary as an instantiation of deliberative virtues: the very structure of judicial procedures hinders our possibilities of testing this intuition, particularly during their decisional phase, a point where we are in no position at all to claim that judges deliberate.

¹¹² *Id.*, § 98.

¹¹³ HABERMAS, *BETWEEN FACTS AND NORMS*, *supra* note 1, at 242, 279–280.

¹¹⁴ Robert Alexy, *Balancing, Constitutional Review, and Representation*, 3(4) INT’L J. CONST. L. 572, 579 (2005).

In order to justify this claim, in what follows I analyze an ideal model of judicial deliberation developed by Conrado Hübner Mendes, as his is the most elaborated scheme in the literature answering what a deliberative constitutional court should look like. I conclude that Mendes's model brings to light limitations courts have in deliberative terms, which are mostly rooted in the anonymity of the discussions taking place in the decisional phase of the judicial process, that is, during the moment in which judges deliberate among themselves before they reach a decision.

My analysis generalizes conclusions drawn by Gastil in his study on the US Supreme Court.¹¹⁵ There, he argues that the judiciary "is not an easy governmental body to study, as justices do not generally discuss cases with one another—even when they sit on the same court—and when they do, they do not generally share their discussion with the public."¹¹⁶ Consider, for example, his analysis of the Woodward and Armstrong study,¹¹⁷ which he describes as "[t]he clearest window into the [US] court's functioning." These journalists had access to examine how the court worked internally from 1969 to 1975, an access that "was rare indeed," for "since the book's publication, the Court has allowed no other observer a similar vantage point from which to write."¹¹⁸ The report, Gastil writes, "reminds us that justices are human,"¹¹⁹ especially when considering Woodward and Armstrong's recount of the summer of 1975, a moment in which the Court reached a "low point" from a deliberative perspective: "The net result of this constellation of justices was a string of decidedly nondeliberative cert and case conference meetings."¹²⁰

Gastil concludes that this "is not to say that the Court routinely fails to meet a high standard of deliberation." For him, Woodward and Armstrong's report "does illustrate, however, the susceptibility of the Court to the same distortions, distractions and failures that other deliberative bodies experience."¹²¹

This conclusion is true, but not enough. Courts may or may not be susceptible to deliberative failures, just like any other deliberative agent; that is, I think, obvious. My conclusion is stronger: we are in no position to know whether judges are deliberative during the parts of the procedure in which they gather to decide. Gastil intuitively says this when he says that "[w]ithout a record of the Court's deliberation, one can only judge its deliberation by the legal quality of its judgments, but that, alas, is a task beyond the scope of this book."¹²² His intuition, however, was not taken to its fullest extent, perhaps because he only discusses the US Supreme Court and things may be different in other countries. Yet, to my knowledge, and with one recent exception,¹²³ there are

¹¹⁵ JOHN GASTIL, *POLITICAL COMMUNICATION AND DELIBERATION* (2008).

¹¹⁶ *Id.* at 142.

¹¹⁷ BOB WOODWARD & SCOTT ARMSTRONG, *THE BRETHREN: INSIDE THE SUPREME COURT* (1979).

¹¹⁸ Gastil, *supra* note 115, at 143.

¹¹⁹ *Id.* at 144.

¹²⁰ *Id.*

¹²¹ *Id.* at 145.

¹²² *Id.*

¹²³ ANDRÉ RUFINO DO VALE, *LA DELIBERACIÓN EN LOS TRIBUNALES CONSTITUCIONALES: UN ESTUDIO EMPÍRICO DE LAS PRÁCTICAS DELIBERATIVAS DEL TRIBUNAL CONSTITUCIONAL DE ESPAÑA Y DEL SUPREMO TRIBUNAL FEDERAL DE BRASIL* (2017).

no studies similar to Woodward and Armstrong's, providing us with windows into the internal functioning of other courts in other parts of the world. Besides, their study did not intend to measure deliberation quality and/or quantity by using any instruments similar to those examined in Section 2.

Although the literature on these matters is scarce, Mendes has provided a unique account of what a deliberative court may look like. He starts from the assumption that if "deliberation enhances the condition of constitutional courts, such courts need to be more than 'exemplars of public reason' or 'forums of principle', more than reason-givers or interlocutors. These expressions, and the respective expectations that they convey, are still superficial. They lack more teeth."¹²⁴

Mendes elaborates a scheme of judicial deliberative performance that takes the opacity of the decisional part of the process as a given. This allows me to take his account as a viewpoint from which I can make my own critical remarks. My comments are not meant to debunk Mendes's project, nor I am implying that his effort is a misleading one, or that he is embracing some form of judicial supremacy,¹²⁵ as he is clear that his attempt is to improve the deliberative conditions of a court, not to show that courts are in fact the best candidates to embody the ideals of deliberative democracy. Our projects are different. Rather, I use his model to suggest that even in an idealized form, courts do not successfully meet basic deliberative conditions.

The model is structured around three tiers: the "core meaning," the "facilitators," and the "hedges" of deliberative performance.¹²⁶ They interrelate in order to determine the central deliberative values a court should pursue, and answer questions like who deliberates, who among those deliberating has the power to make decisions, and what the stages of the process itself are.

Mendes distinguishes "three moments in which [judicial] performance might be discerned and appraised, three slices of an overall enterprise":¹²⁷ a pre-decisional, a decisional, and a post-decisional stage. The pre-decisional phase is charged with the task of "public contestation," the decisional with "collegial engagement," and the post-decisional phase with delivering a "deliberative written decision."¹²⁸ He is adamant that courts "may be deliberative in one [of these stages], but not in the other."¹²⁹

As to who deliberates, there are two types of deliberators: judges and interlocutors, namely, a community comprising "all social actors that, formally or informally, address public arguments to the court and express public positions as to the cases being decided. . . . They can influence and persuade, but they cannot decide."¹³⁰ Furthermore, deliberators can be formal, if they are "qualified and entitled to participate of the specific constitutional case (litigants, *amici curiae*, etc.)," or informal, i.e.,

¹²⁴ CONRADO HÜBNER MENDES, CONSTITUTIONAL COURTS AND DELIBERATIVE DEMOCRACY § 100 (2013).

¹²⁵ Walter Sinnott-Armstrong, *Weak and Strong Judicial Review*, 22 L. & PHIL. 381, 381 (2003); Donald Bello Hutt, *Against Judicial Supremacy*, 31 REVUS 7 (2017).

¹²⁶ Mendes, *supra* note 124, at 103.

¹²⁷ *Id.* at 105.

¹²⁸ *Id.*

¹²⁹ *Id.*

¹³⁰ *Id.* at 106.

Table 1. Potential deliberative qualities of constitutional courts in Mendes’s model¹³⁵

Public contestation pre-decisional phase	Collegial engagement decisional phase	Deliberative written decision post-decisional phase
Weak epistemic —	Strong epistemic Communitarian (internal)	Strong epistemic Communitarian (external)
Strong psychological	Weak psychological	Strong psychological
Strong educative	Weak educative	Strong educative
Strong intrinsic	Weak intrinsic	Strong intrinsic

those who “in the attempt to exert an indirect influence on the court, engage in the debates through the various communicative media of the public sphere.”¹³¹

In a nutshell, Mendes’s idea of a deliberative court entails tribunals that foster different deliberative values at different times. Deliberation cuts through the continuum of the trial in different ways depending on the institutional capacity the court has during each procedural moment. It thus promotes public contestation between parties at the early stages of the process, seeking to gather as much information as possible, to include all relevant arguments so that the pool of information is as varied and respectful of the parties as possible. Members of the court should then retire to deliberate behind closed doors. During this moment, collegial engagement is the guiding principle among judges who are to discuss the merits of the information and the arguments made by the interlocutors, weigh those arguments against the law governing the case *sub lite*, according to their principles, and interpretative strategies.¹³² Finally, they should deliver their decision by drafting a text that is the result of the preceding stages, showing that the parties have been considered, their inputs and arguments balanced, and their interests weighed. The decision should be the reflection of the collegiality that drove judges to agree on that particular decision and not on a different one.¹³³ This is a rough image of what Mendes thinks a deliberative constitutional court may look like.

The model does not depend on actual cases or empirical assessments of judicial procedures. What it does instead, is to provide a scheme indicating what we can expect from the court during the procedure, in terms of the values each stage is capable of fostering. He then applies four normative categories of deliberative quality to the pre-decisional, decisional, and post-decisional stages of the judicial process: epistemic, communitarian, psychological, and educative. The interplay between phases and categories results in a scheme that “enumerates the potential deliberative qualities of the three phases,” underscoring “the different ways and degrees in which the values and promises of deliberation are in play”¹³⁴ (see Table 1).

¹³¹ *Id.*
¹³² *Id.* at 108–109, 113–118.
¹³³ *Id.* at 109–113.
¹³⁴ *Id.* at 114.
¹³⁵ *Cf. id.*

Mendes avers that these phases “should not be taken to suggest a discontinuous process with tight and segmented characteristics.”¹³⁶ To put it in the language of deliberative systems theory, the deliberative tasks are distributed. Different agents have different assignments and hold to different deliberative standards.¹³⁷ Still, the model allows to show some difficulties affecting accounts relating deliberative democracy to the judiciary. In particular, that the qualities expected from the court during the decisional phase are not well justified from a deliberative perspective. I will briefly explain each deliberative goal, and then close the section justifying this claim.

The epistemic goal gains strength as the procedure moves forward. The pre-decisional phase contributes “to the multiplication of points of view on a certain controversy [, and it can] at the very least, be a strong practice of information gathering.”¹³⁸ The epistemic goal is weak in this stage because it serves the function of providing interlocutors with a hearing to make their arguments and justify their preferences. It is the court that receives such information, which will be used during the decisional phase, and translated into a decision during the post-decisional phase, which explains the strength of the epistemic objective expected at these two last phases. The learning capacity of the parties is then limited during the first stage, in the sense that it is not a feature that is generally expected from them to develop.¹³⁹ Conversely, given that we assume collegial engagement from judges, “[p]remise unveiling, creativity-sparkling, and truth seeking” are more likely to spring during the decisional phase. The post-decisional stage has the epistemic function of “supplementing the next cases with densely drafted precedents,”¹⁴⁰ in order to avoid future cases from having “to re-inaugurate the deliberative chain from scratch, wasting the argumentative accomplishments and progress of previous cases.”¹⁴¹

The communitarian goal refers to the capacity of each stage to achieve consensus. This goal is absent during the pre-decisional stage, as “it cannot properly have a special commitment to reduce disagreement.” The decisional stage, on the other hand, “has the responsibility of constructing an institutional and de-personified decision,”¹⁴² which should be accepted by the parties at the post-decisional stage.

The psychological goal refers to the “sense of respect instilled among the participants of deliberation.”¹⁴³ For Mendes, the court shows interlocutors respect through a “genuinely porous public contestation and a carefully drafted decision.”¹⁴⁴ Also, collegiality demands intra-institutional respect for the court to maintain “its capacity of deliberation.”¹⁴⁵

¹³⁶ *Id.*

¹³⁷ Robert Goodin, *Sequencing Deliberative Moments*, 40(2) ACTA POL. 182, 189 (2005); ROBERT GOODIN, INNOVATING DEMOCRACY. DEMOCRATIC THEORY AND PRACTICE AFTER THE DELIBERATIVE TURN § 186 (2008); Carolyne Hendriks, *Integrative Deliberation: Reconciling Civil Society’s Dual Role in Deliberative Democracy*, 54 POL. STUD. 486, 499 (2006).

¹³⁸ Mendes, *supra* note 124, at 114.

¹³⁹ *Id.* at 115.

¹⁴⁰ *Id.* at 114.

¹⁴¹ *Id.*

¹⁴² *Id.* at 115.

¹⁴³ *Id.*

¹⁴⁴ *Id.*

¹⁴⁵ *Id.*

A deliberative process should also educate participants, particularly, at the first and the last stages. The decisional stage, however, is not expected to be a moment during which judges “who deliberate on a routine basis” may learn too much; skill enhancement would be a frivolous expectation,” notwithstanding they may still “refine their knowledge on the respective topic.”¹⁴⁶

All these considerations are instrumental. Yet, in order to make his case stronger, Mendes offers a set of non-instrumental reasons why deliberation may be a desirable thing in the courtroom. During the pre-deliberative stage, there are intrinsic reasons to have a deliberative procedure based on the opportunity they give individuals to “make themselves heard and realize their argumentative autonomy before the court.”¹⁴⁷ The same happens with the post-decisional stage, insofar as a deliberative written decision gives the court the “chance to act as a catalyst of external deliberation and to work as an open and accessible ‘forum of contestation.’”¹⁴⁸ Things are different at the decisional stage, for courts may or may not be deliberative in this part of the process. So, if judges promote the sense of respect and recognition each individual deserves through public contestation and through the deliberative written decision, then the court will have performed its tasks. Though present,¹⁴⁹ intrinsic reasons for deliberation wane once judges gather to debate behind closed doors.

Critical remarks are in order. From the preceding description, it becomes clear that courts experience non-negligible limitations as deliberative bodies in the following senses. First, there is a disconnection between the epistemic and the psychological goals, for the fact that the model considers interlocutors as providers of information but not as active parts of the process of deliberation itself evinces the passive role they have in decisional and deliberative terms. The respect shown to interlocutors is not manifested by acknowledging their capacity to discuss and decide but to raise arguments and preferences, which remain apparently fixed during the first stages of the procedure. There is no reason to think that the judicial process would serve to change the point of view of the parties at the trial, or that the decision promotes anything different from mere acceptance, let alone that it will allow one party to share or somehow identify with the views of its counterparts. Moreover, the epistemic strength of the post-decisional phase is weighed vis-à-vis other courts, and not with respect to the parties.

Hence, the epistemic and the psychological goals show that the more the procedure moves from the pre-decisional to the post-decisional phase, the less room there is for the interlocutors to participate. This generates a tension with the notion of deliberative democracy that every person potentially affected by a decision should be included in a more substantive way than by mere aggregation. Mendes’s model also sees the epistemic goal of deliberative democracy in a narrower way than the traditional understanding of the epistemic values given to deliberative processes, which, among

¹⁴⁶ *Id.* at 113.

¹⁴⁷ *Id.* at 116. Likewise, ALON HAREL, *WHY LAW MATTERS* § 192 (2014).

¹⁴⁸ Mendes, *supra* note 124, at 116.

¹⁴⁹ *Id.* at 117.

others, includes the increase of the relevant information available for participants,¹⁵⁰ the detection of factual and logical mistakes,¹⁵¹ the control of irrational factors and the filtering of irrational preferences,¹⁵² and the avoidance of inequalities of information and manipulation of the political agenda.¹⁵³

None of these epistemic functions is fully met in the judicial model portrayed by Mendes, because the only instance in which participants have the chance to be active, the pre-decisional phase, is more aggregative than deliberative. There, preferences remain fixed, learning is limited, and participation adversarial. An example of this lack of deliberation between parties is the absence of communitarian goals at the pre-decisional phase; parties are not expected to reach agreements, neither consensual nor incompletely theorized.¹⁵⁴

The lack of fit of a constitutional court to an ideal of deliberative democracy is also demonstrated by the portrayal of the educative function. Parties are expected to learn and to provide the court with information at the pre-decisional level. But to conclude, from this, that the educative function is likely to be strong at the pre-decisional phase is, I believe, expecting too much. We must bear in mind that trials are adversarial procedures: parties learn from their counterparts, not to find common ground with opposite sides but with the aim of rebutting their arguments. Parties do not attempt to convince their counterparts when they go to court, but to convince judges that it is their argument the one that should prevail. Calling this learning and education seems like an overstatement.

Mendes's scheme also shows that the obscurity of the decisional phase works against the deliberative quality of the procedure. A close look to the goals deliberative courts are expected to achieve shows that in those moments where we have disclosed access to the arguments and discussions raised during the process, where we can test and measure the quality and quantity of the arguments wielded, that is, the pre and the post-decisional phases, the participation of the interlocutors is not fully deliberative: parties are either sources of inputs in an adversarial, non-dialogical sense, as they do not discuss with the expectation of changing each other's convictions or preferences, or they are passive addressees of the court's decision.

These problems are starker within the decisional phase. Here, the goals we expect the court to achieve become less observable, to the extent judges discuss the merits of the interlocutors' arguments behind closed doors. We expect that there will be a strong epistemic thrust during the decisional phase resulting from the collegiality that is to inspire the courts' proceedings; we expect judges to believe in "a supra-individual good that they can reach together, and on which the external respectability of their decision will depend."¹⁵⁵ Yet, we have no way of testing whether this common purpose

¹⁵⁰ Bernard Manin, *On Legitimacy and Political Deliberation*, 15(3) POL. THEORY 338, 349 (1987); CARLOS S. NINO, THE CONSTITUTION OF DELIBERATIVE DEMOCRACY §§ 117–128 (1996); Martí, *supra* note 1.

¹⁵¹ Nino, *supra* note 150, at 124; Martí, *supra* note 1.

¹⁵² Martí, *supra* note 1.

¹⁵³ *Id.* at 43.

¹⁵⁴ On the notion of incomplete theory agreements, *see, e.g.*, CASS SUNSTEIN, LEGAL REASONING AND POLITICAL CONFLICT §§ 35–60 (2018).

¹⁵⁵ Mendes, *supra* note 124, at 134.

is what actually inspire judges while they decide. Conversely, what we do know is that ideology can move judges in one direction or another,¹⁵⁶ and that their fidelity to the legal framework and the facts of the case are not to be taken for granted.¹⁵⁷

We expect the judges' discussions to be infused with cognitive modesty—"a logical and moral condition of preference transformation,"¹⁵⁸ exhorting judges to "make themselves vulnerable to the scrutiny of their fellow colleagues."¹⁵⁹ But, again, we do not have a way to test the extent to which this happens. Moreover, we expect cognitive ambition to "fuel collegiality with an investigative energy without which the epistemic promise of deliberation gets anemic and fatigued."¹⁶⁰ But again, how do we know that judges are not "'advocates of a position' but 'students of an issue' . . . relentless in the search of the best decision?"¹⁶¹ We do not have the tools to know that because we are not allowed to check whether this is the case.

Finally, we expect empathy, that is, that courts have "the ability of vicariously [imagine] the points of view that were not formally voiced in the course of the judicial process."¹⁶² For Mendes, this is the "principal corrective a constitutional court can have against a poorly pre-decisional phase."¹⁶³ Judges should then be able to "go beyond the arguments [they were] able to collect in the pre-decisional phase through empathetic imagination of the potential community of interlocutors."¹⁶⁴ Yet, our hopes that courts will put themselves in the shoes of others are not supported by our knowledge of the social gaps existing between judges and ordinary citizens. There is no guarantee, no matter how heterogeneous the court is in terms of its composition, that its members will be able to know what are the preferences, lacks, needs, arguments, etc., of a community that is more diverse than the court can ever be. Even if, as Ely argued, judges agreed to protect the interests of the minorities who have not been able to reach the court,¹⁶⁵ it remains difficult to imagine judges being capable of performing this representative task—the fact that constitutional courts are counter-majoritarian, does not mean that they are pro-minoritarian. Even if *arguendo* one assumed that this is indeed the role of a court, there is still no correlation between the duty we expect judges to fulfill and their actual capacity to do so.¹⁶⁶

¹⁵⁶ LAWRENCE TRIBE, *GOD SAVE THIS HONORABLE COURT: HOW THE CHOICES OF SUPREME COURT JUSTICES SHAPES OUR HISTORY* (1985).

¹⁵⁷ JEFFREY SEGAL & HAROLD SPAETH, *THE SUPREME COURTS AND THE ATTITUDINAL MODEL REVISITED* § 53 (2002); Jeffrey Goldsworthy, *The Limits of Judicial Fidelity to Law: The Coxford Lecture*, 24(2) CANADIAN J.L. & JURISPRUDENCE 305 (2011).

¹⁵⁸ Mendes, *supra* note 124, at 134.

¹⁵⁹ *Id.*

¹⁶⁰ *Id.* at 135.

¹⁶¹ *Id.*

¹⁶² *Id.*

¹⁶³ *Id.*

¹⁶⁴ *Id.* at 136.

¹⁶⁵ JOHN ELY, *DEMOCRACY AND DISTRUST: A THEORY OF JUDICIAL REVIEW* § 151 (1980).

¹⁶⁶ ROBERTO GARGARELLA, *LA JUSTICIA FRENTE AL GOBIERNO: SOBRE EL CARÁCTER CONTRAMAYORITARIO DEL PODER JUDICIAL* § 181 (1996).

The anonymity of the decisional phase is at the core of these problems. We have the tendency to expect many things from courts,¹⁶⁷ but when we ask whether they are capable to live up those expectations, our best answer has to be found in non-procedural reasons and focus on their outcomes. But then, once we make that move toward outcomes, we tend to leave procedural reasons aside. Moreover, assuming that they can, and that we can envisage schemes that measure the quality and quantity of judicial deliberation, we still lack the access to the fora where that deliberation should occur.

4. Conclusion

Despite the usefulness of the instruments here examined, this article shows the areas in which measuring deliberation is still a work in progress. Problems with the normative foundations of the coding schemes, the use authors make of key theoretical concepts, and their choice for not measuring sincerity weakens the defense of their employment as justified measuring instruments of deliberation.

The article also showed that despite the variety of settings the coding schemes have been applied to, courts are not among them, and that in order to make sense of this omission we have to look at the way judicial procedures are set up. By focusing on Mendes's model of deliberative constitutional courts, Section 3 further strengthened the claim that, in the current state of the literature in deliberative democracy, there is little reason why one would have a high regard for judges in terms of their deliberative capacities: this could or could not be the case. The fact is not only that we do not know but also that unless we change the way decisional phases of judicial procedures are structured, we are unable to know.

I must be careful not to overstate my conclusion, as our knowledge of the effects of anonymity in political arenas is limited.¹⁶⁸ Also, recent scholarship suggests that the benefits deliberative democrats see in transparency and publicity are contingent. It also indicates that in opaque and anonymous settings, even the most self-interested and market-oriented actors adapt to deliberative norms.¹⁶⁹ This means that it may be perfectly legitimate to keep the decisional phase of judicial procedures anonymous. Nonetheless, this scholarship is still too underdeveloped, and too focused on one specific kind of actor, namely, lobbyists, who operate with incentives which are, *prima facie*, different from the ones judges have in their institutional environment. There are studies that follow a strand initiated by Posner that link judicial performance and incentives, where anonymity and political insulation play an important role.¹⁷⁰ These

¹⁶⁷ JEREMY WALDRON, *THE DIGNITY OF LEGISLATION* § 5 (1999).

¹⁶⁸ JAMES A. GARDNER, *Anonymity and Democratic Citizenship*, 19 WM. & MARY BILL RTS. J. 927, 930 (2011).

¹⁶⁹ DANIEL NAURIN, *DELIBERATION BEHIND CLOSED DOORS* (2007).

¹⁷⁰ RICHARD POSNER, *What do Judges and Justices Maximize? (The Same Thing Everybody Else Does)* 3(1) SUP. CT. ECON. REV. 1 (1993).

studies, however, do not measure the effects of those incentives in terms of the effects that they may have in the *deliberative* performance of judges.¹⁷¹

More research is needed on these matters. For now, we can conclude that the scholarship on present-day deliberative democracy proves that when we ask ourselves how people deliberate, we can mostly answer with regard to non-judicial settings. In this ambit, our answer will have to bear in mind that the current instruments we use to determine how people deliberate compromise theoretical ideals to different degrees. My take on this ideal theory/practice divide is that the compromise is justified only within certain domains, for example, when it comes to measure authenticity. But there is certainly work to be done with regard to the philosophical foundations of the measuring mechanisms. In this respect, scholars need to further think whether Habermas's is still the right approach to follow, and, should they decide to stick to that approach, they should also answer *which* Habermas is grounding their schemes—the complexity of his work demands such specification.

Regarding the *judicial* side of the story, this article shows that given the very structure of the courts' procedures, we know a lot less and that we are hindered from knowing much more.

¹⁷¹ Martin R. Schneider, *Judicial Career Incentives and Court Performance: An Empirical Study of the German Labour Courts of Appeals*, 20(2) EUR. J.L. & ECON. 127 (2005); Alessandro Melcarne & Giovanni Ramello, *Judicial Independence, Judges' Incentives and Efficiency*, 11 (2) REV. L. & ECON. 149 (2015).