The Jean Monnet Center for International and Regional Economic Law & Justice

THE NYU INSTITUTES ON THE PARK

THE JEAN MONNET PROGRAM

J.H.H. Weiler, Director
Gráinne de Burca, Director

Jean Monnet Working Paper 5/19

Johann Laux

Public Epistemic Authority
An Epistemic Framework for the Institutional Legitimacy of International Adjudication

NYU School of Law • New York, NY 10011
The Jean Monnet Working Paper Series can be found at
www.JeanMonnetProgram.org

All rights reserved.
No part of this paper may be reproduced in any form
without permission of the author.

ISSN 2161-0320 (online)
Copy Editor: Danielle Leeds Kim
© Johann Laux 2019
New York University School of Law
New York, NY 10011
USA

Publications in the Series should be cited as:
AUTHOR, TITLE, JEAN MONNET WORKING PAPER NO./YEAR [URL]

Public Epistemic Authority

An Epistemic Framework for the Institutional Legitimacy of International Adjudication

Johann Laux[*]
University of Oxford

Abstract This paper suggests a new framework for the legitimacy of international adjudication relative to other branches and modes of public decision-making. Named 'Public Epistemic Authority', this new framework approaches judicial authority from its often-overlooked institutional dimension. In this dimension, the collective decision-making setting of international courts must find its due consideration. The paper draws on the mechanisms of collective wisdom popular with epistemic theories of democracy to establish a truth-tracking benchmark for international adjudication. Based on this benchmark, the paper suggests an epistemically improved institutional design of international courts and formulates several epistemic principles for judicial reasoning.

[*] Affiliation: University of Oxford, Oxford Internet Institute; Emile Noël Fellow, New York University, 2019; Research for this paper was made possible by the generous support of the German Academic Exchange Service (DAAD).
Contact Email: johann.laux@oii.ox.ac.uk

I. INTRODUCTION

Scholars of global governance have not failed to observe the quantitative increase in international adjudication and the qualitative change in the role of international courts and tribunals (ICs) from settling disputes towards solving problems for a global society.[1] They speak of the international "public authority" which ICs exercise.[2] Scholars have further noticed a growing plurality and diversity of international institutions and regimes which compete to address similar, sometimes overlapping global issues.[3] The questions of legitimacy this development raises have also not been overlooked.[4] The legitimacy of the authority of ICs, however, is usually studied only within two of the three dimensions of the concept of legitimacy. Scholars problematize the *moral* dimension of legitimacy and present instrumental, procedural, or other accounts which rival for the best justification of judicial authority and its alleged right to rule.[5] They also study the

---

[1] Since 2002, international courts have rendered more judicial decisions every single year than was the case from time immemorial up to 1989, cf. Armin von Bogdandy and Ingo Venzke, *In Whose Name? A Public Law Theory of International Adjudication* (Oxford: Oxford University Press, 2014), 1.

[2] Ibid. See also: Cesare P.R. Romano, Karen J. Alter, and Yuval Shany (eds.), *The Oxford Handbook of International Adjudication* (Oxford: Oxford University Press, 2014); Benedict Kingsbury, "International Courts: Uneven Judicialization in Global Order", in: James Crawford and Martti Koskenniemi (eds.), *Cambridge Companion to International Law* (Cambridge: Cambridge University Press, 2012), 202-228; Geir Ulfstein, "The International Judiciary", in: Jan Klabbers, Anne Peters, and Geir Ulfstein (eds.), *The Constitutionalization of International Law* (Oxford: Oxford University Press, 2009), 126-152; Matthias Goldmann, *Internationale Öffentliche Gewalt: Handlungsformen internationaler Institutionen im Zeitalter der Globalisierung* [International Public Authority: International Institutions and Their Instruments in the Age of Globalization] (Berlin: Springer, 2015).

[3] Cf. Gráinne de Búrca, "Contested or Competitive Multilateralism: A Reply to Julia C. Morse and Robert O. Keohane", in: (2016) 5 *Global Constitutionalism* 3, 320–326. See further: Julia C. Morse and Robert O. Keohane, "Contested Multilateralism", in: (2014) 9 *Review of International Organizations* 4, 385–412; Nico Krisch, *Beyond Constitutionalism: The Pluralist Structure of Postnational Law* (Oxford: Oxford University Press, 2010).

[4] See, amongst others: Wojciech Sadurski, "Supranational Public Reason: On Legitimacy of Supranational Norm-Producing Authorities", in: (2015) 4 *Global Constitutionalism* 3, 396–427; Nienke Grossman, "The Normative Legitimacy of International Courts", in: (2013) 86 *Temple Law Review* 1, 61–106; Nienke Grossman, Harlan Grant Cohen, Andreas Follesdal and Geir Ulfstein (eds.), *Legitimacy and International Courts* (Cambridge: Cambridge University Press, 2018); Karen Alter, Laurence Helfer and Mikael Madsen, *International Court Authority* (Oxford: Oxford University Press, 2018); Robert Howse, Hélène Ruiz-Fabri, Geir Ulfstein and Michelle Q. Zang (eds.), *The Legitimacy of International Trade Courts and Tribunals* (Cambridge: Cambridge University Press, 2018); Cecilia M. Bailliet and Nobuo Hayashi (eds.), *The Legitimacy of International Criminal Tribunals* (Cambridge: Cambridge University Press, 2017).

[5] Instead of providing an exhaustive list, just confer the following papers: for an instrumentalist account, see Grossman, "The Normative Legitimacy of International Courts", op. cit. 4; for a proceduralist account, cf. von Bogdandy and Venzke, op. cit. 1; for an interactional account, cf. Jutta Brunnée and Stephen J. Toope, *Legitimacy and Legality in International Law: An Interactional Account* (Cambridge: Cambridge University Press, 2010); for an example of a mixed account, see also: André Nollkaemper, *International*

empirical dimension of legitimacy and analyze what people believe to be the legitimate exercise of judicial authority.[6] Legitimacy, however, also has an *institutional* dimension, here assessing the relative competence of the judiciary. Though the institutional and the moral dimension overlap, they are not identical, this paper argues.

Moving from safeguarding peace amongst nations to solving collective problems raises the demands for the institutional competence of ICs. Especially where legal regimes overlap, the problem of what kind of institution attempts to solve what kind of problem is only inadequately addressed by the moral and the empirical dimensions of legitimacy. There, the democratic legitimacy of authorities beyond the state remains a major concern, often barring more granular institutional questions from view. This might be one reason for why the debate about the proper institutionalization of judicial authority is only begging to take hold amongst theorists.[7]

It is of course not the case that the broader literature on courts has disregarded problems of judicial competence. Both the national as well as the international judiciary have been scrutinized to varying degrees, often along similar lines. For example, the ability of international judges to handle scientific evidence and expertise is being questioned,[8] and such is that of their national counterparts.[9] Other relevant themes are the adequate selection process for international judges,[10] the proper decision-making protocol for judges on a panel,[11] and the allocation of law-making authority between

---

Adjudication of Global Public Goods: The Intersection of Substance and Procedure, (2012) 23 *European Journal of International Law* 3, 769–791.

[6] See, for example: Mark A. Pollack, "The Legitimacy of the European Court of Justice: Normative Debates and Empirical Evidence", in: Grossman et. al., op. cit. 4, 143–173; Andrea K. Bjorklund, "The Legitimacy of the International Centre for Settlement of Investment Disputes", in: *ibid.*, 234–283.

[7] Making this claim for the case of judicial review by (national) constitutional courts: Mattias Kumm, "On the Representativeness of Constitutional Courts: How to Strengthen the Legitimacy of Rights Adjudicating Courts without Undermining Their Independence", in: Christine Landfried (ed.), *Judicial Power: How Constitutional Courts Affect Political Transformations* (Cambridge: Cambridge University Press, 2019), 281–291, 283. ICs do not seem to be different in this regard.

[8] Makane Moïse Mbengue, "International Courts and Tribunals as Fact-Finders: The Case of Scientific Fact-Finding in International Adjudication", in: (2011) 34 *Loyola of Los Angeles International and Comparative Law Review* 1, 53–80.

[9] Scott Brewer, "Scientific Expert Testimony and Intellectual Due Process", in: (1997–98) 107 *Yale Law Journal* 6, 1535–1681.

[10] Cf. Mitchel Lasser, "European Judicial Appointments: A Neo-Institutionalist Approach", in: Christine Landfried (ed.), *Judicial Power: How Constitutional Courts Affect Political Transformations* (Cambridge: Cambridge University Press, 2019), 82–108.

[11] Lewis A. Kornhauser, "Deciding Together", in: (2015) 1 *Journal of Institutional Studies* 1, 38-61, 45.

judicial and non-judicial institutions.[12] The list is by no means exhaustive. The problem is rather that the question of judicial competence is usually not conceptualized as part of a distinct dimension of legitimacy. This conceptual neglect has left the debates on the normative grounds of judicial authority and on its institutionalization somewhat disconnected. On one hand, it is left to strong instrumentalist stances to present moral reasons for the desirability of measuring judicial legitimacy according to the outcomes it generates.[13] As a result, genuinely institutional questions become entangled in deep moral controversies such as those between instrumentalists and proceduralists. However, at least with regards to judicial authority it seems obvious that neither pure proceduralism nor pure instrumentalism can hold.[14] On the other hand, those theorists working on institutional questions of judicial authority tend to bracket the problem of fundamental disagreement in society.[15] Consequently, the concept of legitimacy ends up being bracketed as well.

The neglect of the institutional dimension is the first of two oddities this paper seeks to address. The second is that until recently,[16] scholars problematizing judicial competence fail to account for the collective decision-making setting of (apex) courts, not to mention ICs.[17] If the institutional dimension of legitimacy is inherently comparative – as I believe it is – then the relevant object of study is not the competence of *the* judge but of the judicial panel or chamber which renders a decision. As the institutional choice to install an IC may well curtail the authority of existing institutions – both judicial and non-judicial, both national and international – we want to know whether we are better off with

---

[12] Adrian Vermeule, *Judging under Uncertainty: An Institutional Theory of Legal Interpretation* (Cambridge, MA: Harvard University Press, 2006), 2; Adrian Vermeule, "Collective Wisdom and Institutional Design", in: Hélène Landemore and John Elster (eds.), *Collective Wisdom: Principles and Mechanisms* (Cambridge: Cambridge University Press, 2012), 338–367.

[13] See, for example, Nienke Grossman, "The Normative Legitimacy of International Courts", in: (2013) 86 *Temple Law Review* 1, 68–72, claiming that if states are better at complying with international law acting on their own and without being subject to an IC's jurisdiction, then it is difficult to justify the latter's authority.

[14] Pure proceduralism has the problem that it offers little safeguards against the possibility of us erring and legitimizing, for example, witch trials. Pure instrumentalism disregards the value pluralism prevailing in our societies.

[15] Vermeule, *Judging under Uncertainty*, op. cit. 12, 2; Cass R. Sunstein, *A Constitution of Many Minds: Why the Founding Document Doesn't Mean What It Meant Before* (Princeton, N.J.: Princeton University Press, 2009), 215.

[16] James R. Dillon, "Expertise on Trial", in: (2018) 19 *Columbia Science and Technology Law Review* 1, 247–312.

[17] Brewer, "Scientific Expert Testimony and Intellectual Due Process", op. cit. 9.

the new international judiciary than we were with our old regime.[18] Potentially, the nationally diverse composition of ICs may have some legitimizing force through an institutional advantage from higher cognitive diversity, stemming from diverse legal points of view of judges who were trained in different legal traditions.

To get this research agenda off the ground, we need new frameworks to assess the institutional legitimacy of judicial authority. One such framework, called Public Epistemic Authority (PEA), is presented here. PEA seeks to establish a cognitive *and* comparative benchmark for institutional decision-making. The focus on the epistemic – understood here as truth-tracking, as will be explained further below – does not mean, however, that all matters within the institutional dimension are epistemic. Think, for example, of the institutionalization of fairness or representativeness.[19]

Still, even an institutional benchmark requires normativity. The institutional dimension of legitimacy cannot fully generate normativity on its own. In other words, we need the moral dimension of legitimacy to provide us with normative reasons to design institutions which generate outcomes which tend to be correct. The most attractive source of normativity for judicial authority, I suppose, is a democratic one. Judicial authority is one type of public authority and arguably, democracy has replaced other forms of justification for public authority such as technocratic expertise, tradition, or any "natural order" of things.[20] The legitimacy of democracy, however, has itself been challenged for some time now by findings of social choice theory. This challenge has been based first and foremost on Arrow's Impossibility Theorem which claims to prove that democratic collective decision-making processes cannot be both fair and rational, as there is no democratic voting procedure available which aggregates existing individual preference rankings into a single, consistent collective outcome.[21] More recently, this first skeptic wave of anticipating social choice theorems within democratic theory has been followed

---

[18] This argument is an application of Raz' Service Conception of Authority, see Joseph Raz, *The Morality of Freedom* (Oxford: Clarendon Press: 1986). For an application of Raz' Conception to the realm of international law, see "The Normative Legitimacy of International Courts", op. cit. 13, 100.

[19] On representativeness, see recently: Matthias Kumm, "On the Representativeness of Constitutional Courts", op. cit. 7.

[20] Richard H. Pildes and Elizabeth S. Anderson, "Slinging Arrows at Democracy: Social Choice Theory, Value Pluralism, and Democratic Politics", in: (1990) 90 *Columbia Law Review* 8, 2121–2214, 2123.

[21] Ibid., 2129; Kenneth Arrow, *Social Choice and Individual Values* (New Haven, CT: Yale University Press, 1951).

by a second, positive wave which builds on the "wisdom" which may emerge in collective decision making. This 'epistemic turn'[22] in democratic theory utilizes mathematical theorems such as the Diversity Trumps Ability Theorem (DTA) and the Condorcet Jury Theorem (CJT) to argue for democratic rule via its inclusion of diverse opinions (DTA) and large groups (CJT). While the first and the second wave disagree in their conclusions about the legitimacy of democratic rule, they are both based on formal models (instead of empirical observations) and share some Condorcetian roots (Arrow's theorem can be understood as a generalization of Condorcet's Paradox).[23] Both waves also produced claims for re-allocating decision-making authority between institutions, though pointing in opposite directions. Under the influence of Arrow's Theorem, legal and political theorists questioned whether legislatures and executives deserve judicial deference.[24] More recently, and on the other side of the spectrum, Adrian Vermeule sought to maximize "the epistemic quality of the laws" and argued for a "large-scale reallocation" of lawmaking authority away from courts, based on the mechanisms of collective wisdom.[25]

These developments must spur the interest of those scholars who are working on the legitimacy of IC authority. The research objective for this paper can thus be stated as follows: It seeks to shed light on the often-neglected institutional dimension of legitimacy and to fill a gap in the literature on IC authority. The paper then begins its inquiry from the assumption that judicial competence is best measured within a collective paradigm of competence, which accounts for the collective decision-making setting of most ICs. Finally, it suggests one particular framework, PEA, as a way to operationalize a collective paradigm of competence in the institutional dimension of IC legitimacy. This framework will be spelled out in terms of its implications for the institutional design of ICs and for the judicial reasoning of international judges. As such, PEA is primarily concerned with an appreciation of the second, positive wave of social choice theory within democratic

---

[22] Hélène Landemore, "Beyond the Fact of Disagreement? The Epistemic Turn in Deliberative Democracy", in: (2017) 31 *Social Epistemology* 3, 277–295.

[23] Pildes and Anderson, "Slinging Arrows at Democracy", op. cit. 20, 2131. While the CJT achieves a positive result, as it shows that simple majority voting can do a good job at producing correct outcomes, given certain conditions, the Paradox attains a negative result, as it shows how pairwise voting over three or more alternatives under majority rule can lead to intransitive (or cyclic) outcomes. Cf. John W. Patty and Elizabeth Maggie Penn, *Social Choice and Legitimacy: The Possibilities of Impossibility* (New York: Cambridge University Press, 2014), 12–13.

[24] See the references in Pildes and Anderson, "Slinging Arrows at Democracy", op. cit. 20, 2125.

[25] Vermeule, "Collective Wisdom and Institutional Design", op. cit. 12., 349.

theory. The only prior exploration of the mechanisms of collective wisdom for the design of courts presented by Vermeule focuses on the U.S. Supreme Court and Congress.[26] Though this paper takes its cue from Vermeule's work, it makes several strong departures from its predecessor.

First, we need to be clear about the normative limits of drawing on theorems such as the DTA and the CJT. Those theorems cannot themselves provide the normativity needed to justify a right to rule for an authority or an obligation to obey for its subjects. Any institutional benefit of the doubt for, say, parliamentary decision making based on ideas of collective wisdom is necessarily contingent on the institutional features prevailing at the time of inquiry. At the same time, it cannot be ruled out that a framework built upon collective wisdom is incompatible with a particular moral conception of legitimacy. I believe that some form of 'epistemic proceduralism'[27] can provide a morally compatible source of normativity for PEA. This argument, however, has to be made elsewhere. While the next section will explicate the source of normativity further, the focus of this paper must remain with the institutional dimension of legitimacy.

The paper proceeds as follows: Section II discusses the institutional dimension of legitimacy from an epistemic lens. It introduces certain competence-based challenges to the legitimacy of judicial authority which occur in international adjudication which need to be addressed. It then presents PEA as one possible framework in which to give an answer to these challenges. Scholars disagreeing with the plausibility of PEA are invited to suggest rivaling frameworks for the institutionalization of judicial authority beyond the state. Section III then applies the framework of PEA. It does so in two different strategies: it discusses both options for institutional (re-)design of ICs and suggestions for principles of institutional judicial reasoning.

Finally, one last word of intellectual caution is due. The nature of this paper is exploratory. At times it proceeds in grand strides and moves past deeply controversial issues, especially in moral theory without taking a well-reasoned, definite stand as to their resolution. For the sake of its argument it also assumes the real-world applicability of the CJT and the DTA, despite much well-founded skepticism (see further below). The goal of

---

[26] Vermeule, "Collective Wisdom and Institutional Design", op. cit. 12.

[27] Cf. David M. Estlund, *Democratic Authority: A Philosophical Framework* (Princeton, NJ: Princeton University Press, 2008), 7–8.

the paper is not to 'prove' the applicability of a theorem, but to spell out a conceptual framework with a theoretically plausible truth-tracking benchmark for institutions, to invite comments on its merits, and to open up a debate about an improved institutionalization of judicial authority beyond the state.

II. ICs and the Institutional Dimension of Legitimacy

This section begins by outlining the epistemic-institutional approach PEA seeks to introduce (1.). It is part of any institutionalist approach to clarify the objective function the institution(s) under review are supposed to fulfil. While it is beyond the ambitions of this paper to present its own theory of adjudication, it suggests a meaningful binary description of the judicial task in international adjudication (2.). Following this, two pressing types of competence-based challenges to judicial authority will be discussed: empirical and normative uncertainty (3.). Although the scope of PEA is not limited to decision making under uncertainty, these challenges appear to be particularly apt to demonstrate the need for research on institutional legitimacy. Finally, the framework of PEA is introduced by, first, fleshing out the relationship between the moral and the institutional dimension of legitimacy and, second, explaining the mechanisms that are supposed to drive the emergence of collective wisdom (4).

1. The Epistemic-Institutional Outlook

Let us begin our exploration with a simple observation. International lawyers are first of all legal experts and so are international judges and arbitrators. As much as ICs exercise public authority, they require legitimation. Institutionally, their expertise must thus account for something, especially when curtailing the authority of parliaments, who are not legitimized by the domain expertise of their members. This is not different for international or national courts. Nor is the need for legitimation a particular feature of the judiciary – other non-parliamentary institutions operating at least also in the political realm such as regulatory agencies and central banks have likewise been found to require legitimation in their tendency to usurp decision-making authority.[28] As much as those

[28] As Samuel Issaccharoff writes:

institutions owe their existence to the alleged expertise of their members, their reliability as decision makers must be taken into an account of their legitimate exercise of public authority.

From this institutional perspective, we are thus asking how reliable judges are as decision makers, how their reliability fares in relation to other institutions, and how this affects their legitimacy. To get this research paradigm started, we will need to describe the objective(s) which international adjudication is supposed to achieve (see next sub-section). At the utmost abstraction, we will understand the role of international courts as to reliably determine correctly whether or not the law has been breached. But for our purposes, reliability requires a benchmark. Since this paper suggests an epistemic framework, it endorses a truth-tracking benchmark. This means understanding adjudication (at least partly) as an epistemic problem, i.e. one in which there is a right (or correct) answer. Usually, the proposition that questions of law have right answers is identified with Ronald Dworkin's position in the debate about the contested nature of law.[29] It also matters for the equally contested nature of adjudication. While the former debate is outside of the scope of this paper, the latter is closely related to our inquiry. The conception of adjudication pursued here understands collegiate ICs as aggregating judgments – not preferences – and hence accuracy, i.e. their reliability to "get the right answer" is the standard against which to measure their performance.[30] As mentioned,

---

> "The loss of state competence in democratic societies pushes courts into areas of state responsibility once properly reserved for the political branches. The line between rights declaration and assumption of responsibility for governmental functions has always been permeable. […] But when the loss of competence is coupled with the discrediting of democratic politics, the temptation for judicial action appears inescapable."*

* Samuel Issaccharoff, "Judicial Review in Troubled Times: Stabilizing Democracy in a Second Best World", New York University School of Law, Public Law & Legal Theory Research Paper Series, Working Paper No. 18-51, December 2018, 39.

[29] Cf. Ronald Dworkin, *Taking Rights Seriously* (Cambridge, MA: Harvard University Press, 1977), 14.

[30] Lewis Kornhauser and Lawrence Sager distinguish two conceptions of adjudication: the rendering of judgment and the rendition of preferences. They then suggest three different models of collegial adjudication, each of which comes with a distinct standard against which judicial performance is measured:

> "(1) If collegial courts aggregate the preferences of the judges, then *authenticity*, the extent to which the court's judgment correctly reflects the preferences of the judges, measures the quality of adjudication. (2) If collegial courts aggregate judgments, then *accuracy*, i.e. their ability to "get the right answer" however one defines the right outcome, is the appropriate criterion. (3) If collegial courts are representative institutions that seek to reach the outcome that the represented body would have reached if they deliberated and voted, then we may identify two evaluative criteria: *fit*, or the tendency to arrive at results that the

this does not mean that the proper institutionalization of judicial authority only requires the satisfaction of epistemic demands.

Anyone suggesting a truth-tracking benchmark for judicial decision-making runs into a particular problem. For many types of judicial decisions, there seems to be no procedure-independent standard of correctness. In order to establish whether or not a particular decision was right or wrong, correct or incorrect, better or worse, we would need another judicial procedure. For example, in absence of new evidence, the guilt or innocence of a defendant can only be established in yet another criminal procedure. One response to this problem is to follow Jon Elster in his "negative institutional design"[31], curtailed to reducing only the worst distorting factors on our public decision making, such as well-researched cognitive biases.

PEA, to the contrary, suggests a 'positive' approach. The idea is not to identify any correct answers in particular – or "the truth" – for international adjudicators. Instead, PEA diverts the mathematical, *a priori* models popular with epistemic democrats as bolstering a moral argument for democratic rule into a benchmark for positive institutional design. *A priori* models are, of course, no guarantee that groups actually behave wisely, a point repeatedly made by critics of epistemic democracy.[32] Yet in the institutional dimension of legitimacy, they can give us an idea of groups' cognitive potential and hence a benchmark for the design of collective institutions. PEA therefore needs two legs to walk on: First, it entails a Comparative Test of Legitimacy based on the mechanisms of collective wisdom, which compares *a priori* the truth-tracking potential of institutions with each other or that of alternative designs of the same institution. Second, it features a Behavioral Test of Legitimacy which screens for cognitive or motivational biases and other distortions. Those two tests are connected by the fact that

---

represented group would have reached; and *reliability*, the absence of bad surprises. In this context, they rely on the Condorcet Jury Theorem."*

* Lewis A. Kornhauser, "The Analysis of Courts in the Economic Analysis of Law", in: Gregory A. Caldeira, R. Daniel Kelemen, and Keith E. Whittington (eds.), *The Oxford Handbook of Law and Politics* (Oxford: Oxford University Press, 2008), 698–710, 705–706, orig. ital.; referring to: Lewis A. Kornhauser and Lawrence G. Sager, "Unpacking the Court", in: (1986) 96 *Yale Law Journal* 1, 82–117.

[31] Jon Elster, *Securities Against Misrule: Juries, Assemblies, Elections* (New York, NY: Cambridge University Press, 2013), 84.

[32] See, exemplarily: Jason Brennan, "How Smart is Democracy? You Can't Answer that Question a Priori", in: (2014) 26 *Critical Review* 1–2, 33-58.

the rather demanding preconditions for mathematical group wisdom to actually emerge in real-world institutions require the members of the group to make their decisions independently. Motivational or cognitive biases can violate the independence condition. What may seem like a crude indicator at first can thus, over time and with further research, turn into a real comparative device for scholars of international adjudication, global governance and constitutionalism. Let us therefore look next at the tasks which ICs fulfil.

2. A Binary Task Description for International Adjudication

International adjudication does not fulfil a neat and singular objective function. Rather, it pursues several objectives. Most scholars agree that international adjudication fulfils a dispute settlement function by defusing threats to peace between states.[33] Others name, for example, law identification,[34] the production of legal knowledge,[35] fact-finding,[36] governance[37] and problem-solving[38]. Armin von Bogdandy and Ingo Venzke describe ICs hence as "multifunctional"[39], fulfilling four primary functions: dispute settlement, the stabilization of normative expectations, law-making, and the control and legitimation of public authority.[40] The authors understand this multifunctionality as an outcome of a historical development of ICs moving away from mere dispute settlement towards solving problems for a global society.[41]

---

[33] Cf. José E. Alvarez, "What are International Judges for? The Main Functions of International Adjudication", in: Cesare P. R. Romano, Karen J. Alter, and Yuval Shany, *The Oxford Handbook of International Adjudication* (Oxford: Oxford University Press, 2013), 159–178, 159.

[34] Cf. Samantha Besson, "Legal Philosophical Issues of International Adjudication: Getting Over the "Amour Impossible" between International Law and Adjudication", in: Romano et. al., *The Oxford Handbook of International Adjudication*, op. cit. 33., 413–434.

[35] Cf. Benedict Kingsbury, "International Courts: Uneven Judicialisation in Global Order", in: James Crawford and Martti Koskenniemi (eds.), *The Cambridge Companion to International Law* (Cambridge: Cambridge University Press, 2012), 203–227.

[36] Cf. Alvarez, "What are International Judges for?", op. cit. 33.

[37] Cf. Kingsbury, "International Courts", op. cit. 35.

[38] Harlan Grant Cohen, Andreas Follesdal, Nienke Grossman, and Geir Ulfstein, "Legitimacy and International Courts – A Framework", in: Grossman et. al., *Legitimacy and International Courts*, op. cit. 4, 1–40, 6.

[39] Von Bogdandy and Venzke, *In Whose Name?*, op. cit. 1, 1.

[40] Ibid., 8.

[41] The authors state that in the since 2002, international courts have rendered more judicial decisions every single year than was the case from time immemorial up to 1989, cf. ibid., 1. It is unlikely that this trend has been reversed in the years since their writing.

The plurality of functions of ICs has led to calls for plurality in both theory and research methodology.[42] The epistemic framework suggested here is a response to these calls. To make a contribution, it must find a plausible generalization of at least a subset of functions international adjudication fulfils. The starting point for PEA is to model ICs as first and foremost aggregating judgments.[43] Empirically, ICs may also aggregate the preferences of the contracting states. However, the widespread staffing of ICs with legal experts as judges speaks to their institutional role as judgment aggregators and hence allows their subsequent modeling as such. At the same time, it is not uncommon for judges to publicly claim that judicial reasoning is an act of (legal) cognition.[44] It serves as a justification for judicial review.[45] If a court's written judgments "are presented as acts of cognition, not of will"[46], then this allegedly differentiates adjudication from political decision making.

We can take our next cue again from von Bogdandy and Venzke who identify the institutional role of ICs despite their multifunctionality as to "determine what the law is and apply it to a concrete case"[47]. This is already a non-trivial, binary task description – a norm is the law or is not the law and a concrete case falls under it or does not fall under it. Taking von Bogdandy's and Venzke's description as a starting point, we will say that ICs' institutional task is to *reliably determine correctly whether or not the law has been breached*. At times, this may include determining what the law is.

The last sentence hints at a complication for our binary description. For some ICs, the identification of the law and/or the interpretation of its meaning is their main business. This must have consequences for their institutional design. In the literature on judgment

---

[42] Cf. Theresa Squatrito, Oran R. Young, Andreas Follesdal, and Geir Ulfstein, "A Framework for Evaluating the Performance of International Courts and Tribunals", in: Theresa Squatrito, Oran R. Young, Andreas Follesdal, and Geir Ulfstein, *The Performance of International Courts and Tribunals* (Cambridge: Cambridge University Press, 2018), 3–35.

[43] In the literature parliaments are often modeled as aggregating preferences and courts as aggregating judgments. However, the institutional design of parliaments, including topic-specific committees and the emergence of evidence-based regulatory law, could argue the case to model parliaments as at least also aggregating judgments.

[44] One judge at the Court of Justice of the European Union argues that his court solely performs cognitive tasks, cf. Thomas von Danwitz, *Funktionsbedingungen der Rechtsprechung des Europäischen Gerichtshofs*, 43 Europarecht 769, 769-70 (2008).

[45] Cf. Dieter Grimm, "What Exactly is Political About Constitutional Adjudication?" in: Landfried, Judicial Power, op. cit. 7, 307–317, 309.

[46] Ibid.

[47] Von Bogdandy and Venzke, *In Whose Name?*, op. cit. 1, 7.

aggregation, a simple model describes a judgment or legal opinion as consisting of three elements: (1) the identification of the relevant legal obligation, (2) the determination of the facts of the case, and (3) the final judgment.[48] Courts will emphasize different elements according to their location in the legal system to which principles of institutional design are corresponding. As Lewis Kornhauser breaks it down for the national level, in a civil case in front of the jury, the institutional aim is easy to establish: the institutional designer wants to maximize the probability of a correct determination of the facts.[49] In a criminal trial, the designer seeks to minimize the probability of a wrongful conviction.[50] Appellate court decision making, however, differs again. As Kornhauser writes, it only "superficially exhibit[s] an analogous simplicity: the court wants to reach the legally correct disposition for the legally correct reasons."[51] The problem, of course, is that "each judge may disagree with her colleagues about what legal rule applies and how to apply it."[52] The differences between the judges are not merely due to different informational states. "Rather, the judges disagree fundamentally about the nature and content of the law."[53]

International adjudication features all of the types discussed by Kornhauser. The International Criminal Court should seek to reduce the rate of wrongful convictions. The International Centre for Settlement of Investment Disputes should seek to maximize the probability of its accuracy in factual matters. The European Court of Human Rights (ECtHR) likely features a high degree of disagreement about the content of (human rights) law. Interestingly, only the instruments of international criminal courts contain detailed provisions concerning the presentation and evaluation of evidence. Most treaties creating ICs are relatively silent on evidentiary matters.[54]

The difficulty with reasonable disagreement about the fundamentals of law illustrates the need to become clear about the envisaged domain for PEA in international

---

[48] Abstracted from Christian List and Ben Polak, "Introduction to Judgment Aggregation", in: (2010) 145 *Journal of Economic Theory* 2, 441–466, 442.
[49] Lewis A. Kornhauser, "Deciding Together", in: (2015) 1 *Journal of Institutional Studies* 1, 38–61, 47.
[50] Ibid., fn. 12.
[51] Ibid., 47.
[52] Ibid.
[53] Ibid., 48.
[54] See: Jeffrey L. Dunoff and Mark A. Pollack, "International Judicial Performance and the Performance of International Courts", in: Squatrito et. al., *The Performance of International Courts and Tribunals*, op. cit. 42, 261–287, 265–275.

adjudication. To 'reliably determine correctly whether or not the law has been breached' could be understood as being a task uniquely suited to factual inquiry. Value judgments – which reflect personal and/or social preferences about which reasonable people can disagree – could be argued to not be subjectable to an epistemic benchmark. There can even be reasonable disagreement about what counts as a value judgment.[55] At the same time, there are concerns about 'inductive risk':[56] An idealization of scientific reasoning might downplay the role of value-judgments which scientists must inevitably make in seemingly neutral factual observations. If we want to include value judgments in the domain for PEA, then we are forced to take a cognitivist stance towards value judgments. While I believe that especially moral and political cognitivism are defensible stances,[57] this paper cannot attempt to defense those positions against non-cognitivist arguments. However, it will briefly outline a conceptual third way, which might mitigate non-cognitivist concerns as to the plausibility of PEA.

We can enlarge our analytical toolkit by a third category of 'fact-sensitive principles'. Originally suggested by G. A. Cohen[58] and adapted for the political domain by Hélène Landemore, these are moral or political principles "whose normative value can be at least partially justified by certain facts when those are verified."[59] At the same time, fact-sensitive principles reflect 'basic values', which themselves do not reflect facts.[60] The standard of fact-sensitive principles, or 'context-dependent principles' as Landemore generalizes them, is thus "a combination of empirical accuracy and coherence with underlying basic values"[61]. (International) Law features fact-sensitive principles. Take the precautionary principles as an example: it reflects the basic value of protecting the public

---

[55] See exemplarily the history of the debate on value neutrality in economics, cf. Philippe Mongin, V"alue Judgments and Value Neutrality in Economics", in: (2006) 73 *Economica* 290, 257–286.

[56] Cf. P. D. Magnus, "What Scientists Know Is Not a Function of What Scientists Know", in: (2013) 80 *Philosophy of Science* 5, 840–849.

[57] Others, of course, also hold moral cognitivism to be possible, just see: Jürgen Habermas, "On the Cognitive Content of Morality", (1996), Proceedings of the Aristotelian Society, Volume XCVI, 335–358; On political cognitivism, see: Hélène Landemore, *Democratic Reason: Politics, Collective Intelligence, and Rule of the Many* (Princeton, NJ: Princeton University Press, 2012), 208–231

[58] G. A. Cohen, "Facts and Principles", in: (2003) 31 *Philosophy & Public Affairs* 3, 211–245.

[59] Landemore, *Democratic Reason*, op. cit. 57, 215.

[60] Cohen, "Facts and Principles", op. cit. 58, 214; Drawing himself on: Amartya Sen, *Collective Choice and Social Welfare* (San Francisco: Holden-Day, 1970), 59.

[61] Landemore, *Democratic Reason*, op. cit., 57, 215.

from harm and actualizes itself by the fact of a lack of extensive scientific knowledge on the issue at hand.

Generally speaking, adjudication follows the sometimes-conflicting pursuits of factual accuracy on the one hand and of values such as fairness, representation, or protection from the abuse of power on the other.[62] For example, in criminal proceedings, convictions require a high accuracy of the factual predicates of criminal liability, whereas acquittals do not require similar accuracy for the factual predicates of innocence.[63] However, even acquittals are not fact-free enterprises. Unless courts resort to basic values, there will always be a necessary degree of factual inquiry involved in the decision making. This degree might vary though and be particularly low in private conflict resolution, where the parties can agree on what the facts of the case are.[64] Therefore, even without taking a cognitivist stance on morality and politics, we can expect a sufficiently large domain for PEA to make a contribution to the theory and methodology of IC scholarship.

3. Challenges to Judicial Competence

In the institutional dimension of legitimacy, challenges to judicial competence must be a major concern. As regards fact-finding, for example, the ability of judges to adjudicate (scientific) facts is severely questioned.[65] It is often argued that scientific fact-finding within adjudication, i.e. the handling of scientific evidence as well as of scientific experts' reports and testimony, is particularly taxing for judges. Usually the authors have the natural sciences in mind. When the natural sciences enter the courtroom, it is said, the 'methods of science' meet the distinct 'methods of law'.[66] Judges are described as not competent enough to assess the scientific evidence in front of them[67] or as being troubled by the uncertainty of scientific 'facts', their probabilistic nature in their quest for

[62] See Mirjan Damaska, "Truth in Adjudication", in: 49 *Hastings Law Journal* 2, 289–308.
[63] Ibid., 305.
[64] Ibid., 304.
[65] Just see exemplarily: Brewer, "Scientific Expert Testimony and Intellectual Due Process", op. cit. 9.
[66] See, for example, Makane Moïse Mbengue, "International Courts and Tribunals as Fact-Finders: The Case of Scientific Fact-Finding in International Adjudication", in: (2011) 34 *Loyola of Los Angeles International and Comparative Law Review* 1, 53–80, 56.
[67] Brewer, "Scientific Expert Testimony and Intellectual Due Process", op. cit. 9.

'veracities'.[68] In the case law, parties have questioned the justiciability of matters of "scientific fact and opinion"[69] before, although to no avail.[70]

From an institutional perspective, however, the handling of natural science by courts is not fundamentally different from the handling of social science, especially models from economics.[71] Some authors thus rightly point to the fact that the social sciences, especially the models of economics, are no less complex, probabilistic, or demanding.[72] Even core legal tasks in adjudication have been said to be overburdened by uncertainty. Vermeule has argued that the historical interpretation of a norm in the U.S.-American Constitution can require an institutionally highly demanding factual judgment as to the original intent of the framers.[73]

Analytically, we can differentiate between two types of epistemic uncertainty featuring in legal reasoning: 'empirical epistemic uncertainty' and 'normative epistemic uncertainty'.[74] While the former concerns fact-finding – such as cases under what is also often called 'scientific uncertainty' – the latter concerns normative reasoning. The following sub-sections will discuss two cases under empirical uncertainty and one under normative uncertainty.

a. Empirical Uncertainty: *Methanex* and *Afton Chemical*

The centrality of evidence and fact-finding to international law is widely recognized in the literature.[75] Surfacing new evidence can render a prior judgment incorrect. We rarely

---

[68] Mbengue, "International Courts and Tribunals as Fact-Finders", op. cit. 66, 58–59.

[69] Southern Bluefin Tuna (N.Z. v. Japan, Austl. v. Japan), 23 R.I.A.A. 1, ¶ 41(c) (Perm. Ct. Arb. 2000).

[70] See further the citation in: Mbengue, "International Courts and Tribunals as Fact-Finders", op. cit. 66, 68–69.

[71] Alvarez, "What are International Judges for?", op. cit. 33.

[72] Cf. ibid., 88.

[73] Vermeule, *Judging under Uncertainty*, op. cit. 12.

[74] This useful analytical differentiation stems from Robert Alexy's theoretical rationalization of judicial balancing, which play no further analytical role in this paper. Cf. Robert Alexy, *A Theory of Constitutional Rights*, translated by Julian Rivers (Oxford: Oxford University Press, 2002), 388–425.

[75] See, for example: Peter Tomka and Vincent-Joël Proulx, "The Evidentiary Practice of the World Court" , in: Juan Carlos Sainz- Borgo (ed.), Liber Amicorum Gudmundur Eiriksson (San José: University for Peace Press, 2017), available under: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2693558 (last accessed on: February 19, 2020); Anna Riddell and Brendan Plant, *Evidence Before the International Court of Justice* (London: British Institute of International and Comparative Law, 2009); Caroline E. Foster, *Science and the Precautionary Principle at International Courts and Tribunals* (Cambridge: Cambridge University Press, 2013).

need to invoke a value-laden argument about justice to establish that. Errors are, of course, not limited to criminal law. For example, the ICJ demonstratively erred before in demarcating land and maritime boundaries.[76]

As fact-finding in international law often involves expert evidence, it abounds with institutional questions. When experts are called upon, the epistemic problem shifts from a first-order assessment of the material evidence by the judges to a judicial assessment of the competence of the expert. Epistemologists call this the layperson-expert problem.[77] Legal scholars have also paid increasing attention to the use of experts by ICs over the past decade.[78] When science enters the courtroom, scientific uncertainty might come with it. Several experts might disagree with each other or experts might not be able to conclude their report with a satisfying degree of probability, at least from an adjudicative point of view.

Let us draw on the *Methanex* case of the North American Free Trade Agreement (NAFTA) Tribunal on Jurisdiction and Merits as an example.[79] The decision has been argued by Makane Moïse Mbengue – who claims the probabilistic nature of science to be repelling to judges – to be a positive example of the handling of science, as

> "the tribunal opted for an active approach in its treatment of scientific facts (i.e., it decided to "summarise the principal findings of fact which [it] has made in regard to the scientific issues relating to MTBE" (methyl tertiary-butyl ether))"[80]

---

[76] Cf. the discussion of the cases of *Qatar v. Bahrain* (1994) and *Cameroon v. Nigeria* (2002) in: Riddell and Plant, *Evidence Before the International Court of Justice*, op. cit. 75, 346–352.

[77] Cf. Melissa Lane, "When the Experts are Uncertain: Scientific Knowledge and the Ethics of Democratic Judgment", in: (2014) 11 *Episteme* 1, 97–118.

[78] Cf. Dunoff and Pollack, "International Judicial Performance and the Performance of International Courts", op. cit. 54, 285; Citing: Daniel Peat, "The Use of Court-Appointed Experts by the International Court of Justice", in: (2014) 84 *British Yearbook of International Law* 1, 271–303; Caroline E. Foster, "New Clothes for the Emperor? Consultation of Experts by the International Court of Justice", in: (2014) 5 *Journal of International Dispute Settlement* 1, 139–173; Anna Riddell, "Scientific Evidence in the International Court of Justice: Problems and Possibilities", in: (2009) 20 *Finnish Yearbook of International Law*, 229–258.

[79] Methanex Corp. v. United States, Final Award of the Tribunal on Jurisdiction and Merits, 44 I.L.M. 1345, 1410-12 pt. III, ch. A, 1-2 (NAFTA Ch. 11 Arb. Trib. 2005).

[80] Mbengue, *International Courts and Tribunals as Fact-Finders*, op. cit. 66, 71.

Not all agree with Mbengue's assessment. Jose Alvarez argues that the Tribunal merely answered a legal question concerning discrimination.[81] At this point, the notion of fact-sensitive principles introduced above helps to sharpen our analysis. Even though Alvarez is likely right about the NAFTA Tribunal not trying to answer a scientific question, the principle of non-discrimination is highly fact-sensitive, at least in the context of this case.

In *Methanex*, the eponymous claimant sought arbitration against the United States of America under Chapter 11 of NAFTA as a Canadian investor. Methanex claimed compensation resulting from the losses caused by the State of California's ban on the sale and use of the gasoline additive MTBE. Methanex was at the time of the proceedings the world's largest producer of methanol, a feedstock for MTBE.[82] It argued that the California ban on MTBE was motivated to favor the United States' ethanol industry.[83] Ethanol features in ETBE (ethyl tertiary-butyl ether), which Methanex deems interchangeable with one another as oxygenates.[84] The tribunal, however, decided that

> "[t]his policy was motivated by the honest belief, held in good faith and on reasonable scientific grounds, that MTBE contaminated groundwater and was difficult and expensive to clean up"[85]

Those scientific grounds were provided for by a scientific report authored by the University of California, mandated for by the California law in question, and addressing, among other things, the question whether or not MTBE is carcinogenic.[86]

Now, as Alvarez argues, the confidence of the Tribunal to tackle scientific fact-finding in the case has much to do with the litigants elevating the importance of the scientific evidence in the case and their extensive presentation of witnesses on these matters. The Tribunal would hence have alienated its litigants if it had ignored the evidence adduced.[87] Moreover, the Tribunal did not even attempt to answer the scientific question about the harmfulness of MTBE. Rather, the scientific evidence was deployed to address "a non-scientific, legal question: namely whether the challenged government policy violated

[81] Alvarez, "What are International Judges for?", op. cit. 33, 95.
[82] *Methanex*, op. cit., 79, Part I, page 1, para 1.
[83] Ibid., Part II, page 9, para 25.
[84] Ibid., part II, para 6.
[85] Ibid., part III, para 102.
[86] Ibid., part II, para 10.
[87] Alvarez, "What are International Judges for?", op. cit. 33, 90.

NAFTA by being discriminatory."[88] The strategic virtue of leaving the scientific question unanswered, according to Alvarez, lies in avoiding an unnecessary risk:

> "Had the Methanex tribunal directly found that MTBE causes cancer in humans, that finding would have been subject to delegitimation over time if subsequent scientific studies proved or suggested the contrary."[89]

This is undoubtedly a good strategic reason to leave the scientific questions underlying the case untouched. However, as the non-discrimination principle also captures indirect, *de facto* discriminatory measures,[90] it falls within the category of fact-sensitive principles. It is thus not the case that the Tribunal could ignore the scientific question of the harmfulness of a substance like MTBE in every non-discrimination procedure it hears. In *Methanex*, however, there had already been an appreciation of the harmfulness of MTBE by the Californian law maker, which had itself drawn on a respectable public source, the University of California. A second opinion on the scientific question thus could have been rejected on institutional grounds, as Section III will indeed argue.

The European case of *Afton Chemical*[91], to which we will turn now, helps to substantiate this view. *Afton Chemical* is also concerned with the risks of using a certain gasoline additive. In the case, a reference for a preliminary ruling, the Court of Justice of the European Union (CJEU) had to decide whether the European Union (EU) legislature – namely the European Parliament and the Council[92] – was justified in limiting the use of methylcyclopentadienyl manganese tricarbonyl (MMT), a metallic additive to petrol in Directive 2009/30/EC.[93] The liquid was suspected of having adverse effects on human health and of causing technical damage to vehicles. Many vehicle manufacturers advised

---

[88] Ibid., 95.

[89] Ibid..

[90] This is a common feature of non-discrimination principles in international economic law, cf. Nicolas F Diebold, "Standards of Non-Discrimination in International Economic Law", in: (2011) 60 *The International and Comparative Law Quarterly* 4, 831–865.

[91] CJEU, Case C-343/09, *Afton Chemical Limited v. Secretary of State for Transport* [2010] ECR I-07027.

[92] Recital 35 of the preamble to Directive 2009/30/EC. Directive 2009/30/EC of the European Parliament and of the Council of 23 April 2009 amending Directive 98/70/EC as regards the specification of petrol, diesel and gas-oil and introducing a mechanism to monitor and reduce greenhouse gas emissions and amending Council Directive 1999/32/EC as regards the specification of fuel used by inland waterway vessels and repealing Directive 93/12/EEC, OJ L 140.

[93] Cf. *Afton Chemical*, op. cit. 91.

against its use.[94] Yet the producer of MMT, Afton Chemical, challenged the factual basis of the European regulation: there was no evidence of the substance being so harmful as to justify such stringent limits. However, an unambiguous test method to assess the risks for health and the environment stemming from MMT was not available. Thus, the limits set by the Directive were of provisional nature and open to revision only by future risk assessments. Once a test methodology was available, the limits could be reduced to zero wherever justified by the risk assessment.[95] In their judgment, the CJEU's judges stressed that no public body or independent entity had undertaken a scientific assessment of the effects of MMT. The "widely disparate conclusions reached [...] are dependent on whether the study relied on was carried out by the motor car industry or by the producers of MMT"[96]. Afton Chemical claimed that the limits on MMT were so stringent that they amounted to a *de facto* ban, violating the principle of proportionality, while having erroneously relied on the precautionary principle.[97]

The CJEU thus ran a conventional proportionality test, which featured all four prongs of the European-style proportionality principle: legitimate aim, suitability, necessity, and the balancing stage.[98] For our purposes, only the final prong is of interest here. The judges ascertained whether the EU legislature "attempted to achieve a degree of balance between [...] the protection of health, environmental protection and consumer protection and [...] the economic interests of traders [...]."[99]

Let us for now ignore the taxing complexity of having to balance four interests against each other. The next sub-section will address this issue. As regards our concern with empirical uncertainty in this sub-section, the CJEU acknowledged the scientific uncertainty[100] underlying the decisional situation and thus resorted to the precautionary principle recognized in EU law[101] which allows to *assume* a heavy interference with a recognized interest *with (near) certainty*.[102] The Court emphasized the seriousness of the

[94] Recital 35 of the preamble to Directive 2009/30/EC.
[95] *Afton Chemical*, op. cit. 91., paras 3–4.
[96] Ibid., para 58.
[97] Ibid., para 43.
[98] Ibid., paras 66–69.
[99] Ibid., para 56.
[100] Ibid., para 68.
[101] Ibid., para. 49; Recital 35 of the Preamble to Directive 2009/20/EC.
[102] Precautionary maxims such as "If in doubt, play safe" and "Better safe than sorry" express the rationale behind the precautionary principle, i.e. to refrain from a cost/benefit analysis in situations of

doubts the EU legislature was justified to hold as to whether MMT was a harmless substance.[103] In face of the absence of reliable scientific data due to a lacking public body scientific assessment of MMT and disparate conclusions in the studies conducted by industry, the Court invoked its "insufficiency, inconclusiveness or imprecision" of data formula, which triggers the precautionary principle. Therefore, the uncertainty towards the alleged risk did not render the measures disproportionate.[104]

As in *Methanex*, the CJEU's judges in *Afton Chemical* were not directly concerned with the objective probability of realizing health and environmental protection. Rather, the CJEU and the Tribunal were concerned with the reliability of the European and the Californian legislatures' assessments.[105] As Section III will argue, we might have good institutional reasons for IC judges to defer to the legislatures' assessments, as the CJEU and the Tribunal could be said to have done.

If IC judges are frequently challenged by cases under empirical (scientific) uncertainty, institutional considerations should factor into an assessment of legitimacy of their judicial authority. This is of course not to say that institutional considerations will be able to explain the outcome of the judicial decision making. The comparison of the two cases above suggests that the credentials of the institutions providing the scientific evidence influence the judicial decision. In *Methanex*, the fact that the scientific report came from the University of California was seen as sufficiently thorough, whereas in *Afton Chemical* industry research was regarded as insufficient. Moreover, whether or not the scientific evidence supports or hinders the furthering of economic liberalization and market integration might be another important factor of influence.

---

uncertainty, cf. Deryck Beyleveld and Roger Brownsword, "Emerging Technologies, Extreme Uncertainty, and the Principle of Rational Precautionary Reasoning", in: (2012) 4 *Law, Innovation and Technology* 1, 35-65, 37.

[103] *Afton Chemical*, op. cit. 91., para 59.

[104] Ibid., paras 60–63.

[105] For a more general account of what kind of assessments are scrutinized in balancing, see: Julian Rivers, "Proportionality, Discretion and the Second Law of Balancing", in: George Pavlakos (ed.), *Law, Rights, and Discourse: The Legal Philosophy of Robert Alexy* (Oxford: Hart Publishing, 2007), 167–188, 182–183.

b. Normative Uncertainty: *Google Spain*

Let us now turn to the normative type of epistemic uncertainty. As mentioned with regards to the balancing in *Afton Chemical*, weighing four normative interests is a highly complex task. A good example of balancing under normative epistemic uncertainty can be found in the creation of a qualified 'right to be forgotten'[106] by the CJEU in *Google Spain*.

The case concerned a request for a preliminary ruling on the interpretation of various Articles in the Data Protection Directive 95/46/EC (now replaced by the EU General Data Protection Regulation 2016/679) and of Articles 7 and 8 of the Charter of Fundamental Rights of the EU (the Charter).[107] The by now well-known factual basis of the case is a complaint by Mr. Costeja Gonzáles lodged with the Spanish Data Protection Agency (AEPD), against La Vanguardia, a Spanish daily newspaper, as well as Google. When someone entered Mr. Costeja González's name in the Google search engine, she would obtain links to two pages of La Vanguardia's newspaper, featuring an announcement mentioning a real-estate auction for the recovery of social security debts of Mr. Costeja González. Those pages dated back to the year 1998.[108] Mr. Costeja González requested that La Vanguardia is required to remove or alter the pages in question so that the personal data relating to him no longer appeared or to use certain tools made available by search engines in order to protect the data. He further requested that Google is required to remove or conceal the personal data relating to him so that they ceased to be included in the search results and no longer appeared in the links to La Vanguardia.[109] The AEPD rejected the complaint related to La Vanguardia but upheld it in so far as it was directed against Google.[110]

---

[106] In January 2012, the European Commission had proposed to the European Parliament a "Right to be forgotten and to erasure", cf. European Commission, *Proposal for a Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation)*, Art. 17, COM (2012) 11 final (January 25, 2012).

[107] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ L 281.

[108] Cf. CJEU Case C-131/12, *Google Spain SL and Google Inc. v. Agencia Española de Protección de Datos (AEPD) and Mario Costeja González* [2014] published in the electronic ERC, para 14.

[109] Ibid., para 15.

[110] Cf., Ibid., paras 16-17.

In its judgment, the CJEU had to address the dynamic development of the internet industry since the entry into force of Directive 95/46/EC. As regards the effects of search engine results on a data subject's fundamental rights to privacy and to the protection of personal data, the CJEU held that:

> "[…] the effect of the interference with those rights of the data subject is heightened on account of the important role played by the internet and search engines in modern society, which render the information contained in such a list of results ubiquitous"[111].

In former times, the availability of truthful information which has been published lawfully may not have been a problem. Now, the negative externalities of this "information overload" required regulation.[112] The CJEU held that data subjects might have a right that links to certain information about her or him be removed from a list of search results based on the data subject's name.[113] The now so-called 'right to be forgotten'[114] was in its essence read into the Directive by the Court via its application of the proportionality principle. The actual balancing by the CJEU went as follows:

> "As the data subject may, in the light of his fundamental rights under Articles 7 and 8 of the Charter, request that the information in question no longer be made available to the general public by its inclusion in such a list of results, it should be held […] that those rights override, as a rule, not only the economic interest of the operator of the search engine but also the interest of the general public in finding that information upon a search relating to the data subject's name. However, that would not be the case if it appeared, for particular reasons, such as the role played by the data subject in public life, that the interference with his fundamental rights is justified by the preponderant interest of the general public in having, on

---

[111] Ibid, para 80.

[112] Filippo Fontanelli, "The Mythology of Proportionality in Judgments of the Court of Justice in the European Union on Internet and Fundamental Rights", in: (2016) 36 *Oxford Journal of Legal Studies* 3, 630–660, 642.

[113] *Google Spain*, op. cit. 108, para 99.

[114] Cf. Article 17 of Regulation 2016/679.

> account of inclusion in the list of results, access to the information in question."[115]

As Filippo Fontanelli correctly observes, four values (or rights/principles) are being thrown into the balancing.[116] First, the subject's fundamental rights under Articles 7 and 8 of the Charter as 'compelling reasons' under Article 14 (a) of the Directive; secondly the economic interests of the search engine operator protected by Articles 15-17 of the Charter; thirdly, the general public's interest to obtain information as guaranteed by Article 11 of the Charter; and fourthly, those "particular reasons" which may generate a "preponderant interest of the general public". The complexity of the task is due to the normative uncertainty as to what the Charter demands in this case.

Cognitively, balancing four rights (or interests) can be a "devilish task"[117]. In the case at hand, the right of Mr. Costeja González prevailed over the economic interests of Google and the public's unrestricted freedom of information. Fontanelli argues that the proportionality test in *Google Spain* cannot be rationally reconstructed in order to support the decision. He holds the argument to not be "falsifiable" because too many variables are mere and unprincipled approximations.[118] Thus, an outcome opposite to the CJEU's decision would be equally plausible, "owing to the magnitude of the burden imposed on Google and the number of people whose access to the relevant information is restricted."[119] Little surprisingly, in the judgment, the judges did not explicate the magnitude they attached to the variables of their proportionality test.

There might very well be a rational reconstruction available for the balancing in *Google Spain*. However, it is not the aim of this paper to provide one. What matters more is that by creating an unprecedented right for data subjects and an unprecedented role for Google to operate as a "quasi-public authority" and arbitrate over individual applications of data removal, the CJEU has made a (policy) decision which was governed by normative

---

[115] *Google Spain*, op. cit. 108, para 97.
[116] Cf. Fontanelli, "The Mythology of Proportionality", op. cit. 112, 643.
[117] Ibid.
[118] Ibid., 644.
[119] Ibid., 644-645. Fontanelly continues to argue that internet cases do not appear manageable through balancing as Pareto-optimization cannot occur due to the "massive inefficiency-creating externality (i.e., they create useless restrictions for users whose action is irrelevant to the regulatory purpose pursued)", cf. ibid., 649.

uncertainty. What was demanded by positive EU law was normatively uncertain and the judges filled the gap using their resulting 'epistemic discretion'.[120]

The judges' decision to grant data subjects anything close to a 'right to be forgotten' includes at least a judgment on the effects of restricting access to this kind of information on the internet. From our epistemic-institutional point of view, the question is again how reliable the judges are in making such a judgment. Especially in situations where the European legislature more or less directly turns the judicial judgment into positive law, as it happened later with *Google Spain* and the GDPR, the legitimacy of the normative judgment includes a comparative view on the relative reliability of the CJEU and the European legislature. Other than in *Afton Chemical* and *Methanex*, there was no legislative first opinion available for the judges to defer to. Rather, they themselves provided a first opinion. However, as Article 7 of the Charter replicates Article 8 of the European Convention on Human Rights (ECHR),[121] another institutional question is whether a more reliable proportionality test could have been available if the CJEU had drawn on the reasoning of the ECtHR. Section III is going to provide an answer to that. But first, we will need to introduce our framework of PEA.

4. Turning to the Framework of PEA

So far, we have seen how both empirical and normative epistemic uncertainty raise institutional questions of reliability. A comparative dimension thus opens up: which institution is supposed to answer what kind of question? How reliable or competent we hold an institution to be quite naturally depends on how we assess reliability or competence. As stated in the Introduction, most accounts of judicial competence follow an individualist paradigm, assessing the competence of a court in terms of the competence of *the* judge. PEA instead builds on a collectivist notion of competence, due to its influence by social choice theory and social epistemology.[122] It is therefore that PEA

[120] Cf. Alexy, *A Theory of Constitutional Rights*, op. cit. 74, 388–425.

[121] Cf. Shazia Choudhry, "Article 7 (Family Life Aspects)", in: Steve Peers, Tamara Hervey, Jeff Kenner, and Angela Ward (eds.), *The EU Charter of Fundamental Rights: A Commentary* (Oxford: Hart, 2014), 183–221, 201.

[122] Cf. Hélène Landemore, "Beyond the Fact of Disagreement? The Epistemic Turn in Deliberative Democracy", in: (2017) 31 *Social Epistemology* 3, 277–295; Allen Buchanan, "Political Liberalism and Social Epistemology", in: (2004) 32 *Philosophy & Public Affairs* 2, 95–130.

is going to be most interesting for the design of collegiate courts. However, an individual judge deciding a case would not render the framework inapplicable. An omniscient Herculean judge would fare well under PEA, even better than most collegiate courts. In real world applications though, it is only when judges decide as groups that the mechanisms of collective wisdom can improve our understanding of how to maximize the epistemic design of courts, including ICs.

The current design of (appellate courts and) ICs, however, is built upon an individualist paradigm of competence. Judges are primarily selected according to their individual competence, not according to their contribution to the aggregated competence of the court as a whole. This individualist paradigm has also influenced arguments about legitimacy. Scott Brewer, for example, recognized an emerging norm of "intellectual due process",[123] resulting in extreme skepticism about judicial competence, which itself can be traced back to the individualistic lens he applies. According to Brewer, epistemic deference to scientific experts generates in non-expert judges and jurors beliefs that are, on average, only accidentally and arbitrarily true at best, but not epistemically justified.[124] Brewer's solution to the problem of judicial competence is a "two hat" solution, requiring the very same judge to have both practical (legal) legitimacy and scientific expertise.[125] The implementation of that solution is potentially easier in the U.S.-American legal system, where law degrees are professional degrees students obtain after their bachelor's. Notwithstanding the likelihood of a science-trained lawyer ascending to the bench, we may ask whether a mere bachelor's degree would be enough for judges to keep up with science. Brewer concedes that he has no answer to this.[126]

The impracticality of his solution does not mean that Brewer has not provided judicial authority with a formidable challenge. However, a more promising outlook on judicial competence than Brewer's lies in adopting a perspective from social epistemology. Social epistemologists comparatively evaluate "how well social institutions facilitate the formation, preservation, and transmission of true beliefs."[127] Theories of epistemic

---

[123] Brewer, "Scientific Expert Testimony and Intellectual Due Process", op. cit 9.
[124] Ibid., 1677.
[125] Ibid., 1681.
[126] Ibid., 1679.
[127] Cf. Buchanan, "Political Liberalism and Social Epistemology", op. cit. 122, 98.

democracy thus often amount to investigations into social epistemology.[128] In our context, this allows to consider the entire institution of the court, and not merely the individual judge, as the epistemic "agent of interest".[129] One account of how individuals can yield group beliefs is through aggregating judgments.[130] As mentioned in the Introduction, in our context this means to model adjudication in collegial ICs as judgment aggregation. This collectivist paradigm then allows to compare institutions – and different designs of the same institution – as to their relative epistemic reliability via truth-tracking ratios. For the most part, however, those ratios will have to be determined *a priori*. This important caveat will be reconsidered in the following sub-section. It explains the epistemic turn in democratic theory and shows how PEA lends its methodology but differs conceptually from it (a.). This will also give us a better understanding of how the institutional dimension of legitimacy is distinguishable from the moral dimension. After this conceptual exploration, two practical implications of the framework are sketched out: recommendations for institutional (re-)design and judicial reasoning (b.). This step will prepare the following section in which we will apply PEA to the design and reasoning of ICs (and the CJEU in particular).

<u>a. Institutional Legitimacy and Epistemic Democracy</u>

Legitimacy is one of several concepts such as 'justice' or 'fairness' with a seemingly endless list of conceptions offered in the literature. In the moral dimension, legitimacy can tell us which authorities have a 'right to rule', whether they are justified in the use of coercion, and whether their subjects have an 'obligation to obey' the authority.[131] Within international law, legitimacy is normatively understood as a right to rule with a

---

[128] Cf. Hélène Landemore, "Collective Wisdom: Old and New", in: Landemore and Elster, *Collective* Wisdom, op. cit. 12, 19, fn. 15.

[129] For such an approach to the problem of epistemic competence, see: Dillon, "Expertise on Trial", op. cit. 16. Dillon draws, amongst other ideas, on Alvin Goldman's "collective doxastic agents", i.e. the properties of groups as epistemic agents. Cf. Alvin Goldman, "A Guide to Social Epistemology", in: Alvin I. Goldman (ed.) *Reliabilism and Contemporary Epistemology* (New York: Oxford University Press, 2012), 221–247; See also: Philip Pettit, "Groups with Minds of Their Own", in: Alvin I. Goldman and Dennis Whitcomb (eds.), *Social Epistemology: Essential Readings* (New York: Oxford University Press, 2011), 242–268.

[130] See generally: Christian List and Philip Pettit, *Group Agency: The Possibility, Design, and Status of Corporate Agents* (Oxford: Oxford University Press, 2011).

[131] Cf. Estlund, Democratic Authority, op. cit. 27, 41

corresponding duty to obey of its subjects. Hence, if an IC is legitimate, it is justified in issuing a judgment (or opinion or decision) the parties of the case have a moral obligation to obey.[132]

If we were operating in the moral dimension, then a purely epistemic account of a right to rule would run into the problem of justifying epistocracy, the rule of the knowers. As mentioned in the Introduction, such justifications of public authority have been replaced by democratic accounts of legitimacy, which honor public authorities' subjects' standing as being free and equal. This does not yet rule out *any* epistemic requirements for the moral legitimacy of public authorities. This paper, however, makes no attempt at solving the moral problems of justifying authority in such instrumental ways. Let it just be mentioned that there are indeed accounts in temporary scholarship which are trying to do just that. For some scholars, Joseph Raz's Service Conception of Authority offers an avenue to legitimize the authority of ICs. Through an application of Raz's Normal Justification Thesis to the realm of international law, ICs' authority appears to be justified whenever states are better able to fulfil their obligations under international law than without adjudication through ICs.[133] The problem with the Normal Justification Thesis is, first, that it was intended by Raz to give an answer to the problem of subjection – i.e., How can it ever be rational for a person to subject its judgment to that of an authority? – and not the problem of justification.[134] Second, as regards the latter problem, the Normal Justification Thesis cannot survive its contact with the 'fact' of reasonable disagreement in society.[135]

Above it was mentioned that a more attractive source of normativity for the epistemic-institutional accounts lies in David Estlund's 'epistemic proceduralism'.[136] According to Estlund, legitimate (state) authority rests on three necessary conditions: first, the justification of the form of government in question must be acceptable to all reasonable or qualified points of view; second, it must have a better than random probability of

---

[132] Cf. also Cohen et. al., "Legitimacy and International Courts", op. cit. 38, 4.

[133] Grossman, "The Normative Legitimacy of International Courts", op. cit. 4.

[134] See Joseph Raz, "The Problem of Authority: Revisiting the Service Conception", in: (2006) 90 *Minnesota Law Review* 4, 1003–1044, 1012; See further: Stephen Perry, "Two Problems of Political Authority", in: (2007) 6 *APA Newsletter on Philosophy and Law* 2, 31–36.

[135] Though Raz would likely disagree, cf.: Joseph Raz, "Disagreement in Politics", in: (1998) 43 *American Journal of Jurisprudence* 1, 25–52.

[136] Estlund, *Democratic Authority*, op. cit. 27.

choosing just policies; and three, it must do so better than other forms of government.[137] While the first condition accounts for the reasonable disagreement of society, the second and the third condition are actual (but weak) epistemic conditions.

All this goes to show that in the moral dimension of legitimacy, there are instrumental reasons to justify authority which can provide epistemic accounts in the institutional dimension with normativity. There is, of course, no perfect mapping between the moral and the institutional dimension. Not every plausible institutional argument will be compatible with just any moral conception of legitimacy. We also have non-epistemic reasons for institutional choices which promote, for example, fairness and representativeness, moral values which are non-instrumental.[138]

Now, epistemic democrats aim to bolster a moral argument for democracy by drawing on theories of judgment aggregation and computational models of cognitive diversity. The emerging phenomenon is that of 'collective wisdom', i.e. a group's epistemic accuracy in judgment, its truth-tracking capacity.[139] The flow of normativity is reversed here: *because* the institution of democratic rule achieves better results than other forms of government, democracy is morally desirable. Recently, Hélène Landemore's *Democratic Reason*[140] has renewed scholarly interest in collective wisdom and has led to significant debate about its merits.[141] Landemore presents an epistemic defense of democracy against dictatorship and oligarchy. The key argument is that

> "democracy, defined as an inclusive decision rule that combines a deliberative and an aggregative phase, is more likely to yield better

---

[137] Ibid., 134.

[138] For a recent account on representation, see: Kumm, "On the Representativeness of Constitutional Courts", op. cit. 7.

[139] This follows Vermeule's 'baseline conception' of wisdom, cf. Vermeule, "Collective Wisdom and Institutional Design", op. cit. 12, 339–340; See also: Christian List and Robert E. Goodin, "Epistemic Democracy: Generalizing the Condorcet Jury Theorem", in: (2001) 9 *Journal of Political Philosophy* 3, 277–306. Vermeule draws here on Elster, who defines judgment as the ability to give all factors their due weight, cf. Jon Elster, *Sour Grapes: Studies in the Subversion of Rationality* (Cambridge: Cambridge University Press, 1983), 16. However, Elster's conception of collective wisdom is far more demanding: it appeals to the idea that the decision maker must be motivated by the common interest, as both rational belief formation and impartial motivation are necessary conditions of collective wisdom, cf. Jon Elster, *Alchemies of the Mind: Rationality and Emotions* (New York: Cambridge University Press, 1999).

[140] Landemore, *Democratic Reason*, op. cit., 57.

[141] Just see the Symposium on Landemore's *Democratic Reason* in (2014) 26 *Critical Review* 1–2.

> solutions and predictions on political questions than less inclusive decision rules."[142]

The 'proof' of democracy yielding better results than other decision rules is provided by the *a priori* mechanisms of collective wisdom. The contributions in the field of epistemic democracy rely on formal models with varying degrees of emphasis: Landemore's *Democratic Reason* first and foremost on Lu Hong's and Scott Page's Diversity Trumps Ability Theorem (DTA), Robert Goodin's and Kai Spiekermann's *Epistemic Democracy* further develops and refines the Condorcet Jury Theorem (CJT).[143] For legal decision making, Cass Sunstein as well as Vermeule have discussed evolutionary aggregation through precedent.[144] In a nutshell, the DTA suggests that in problem-solving, a randomly selected collection of (sufficiently intelligent) problem solvers outperforms a collection of the best individual problem solvers.[145] In its classic formulation, the CJT states that when a group makes a choice between two alternatives, one of which is correct, and if the members of the group individually are more likely than not to identify the correct alternative, then the probability of the majority of the group to vote for the correct alternative approaches certainty as the number of voters in the group goes to infinity.[146]

The appeal of those theorems to democratic theorists is easy to identify: Democratic rule is supposed to include large groups (CJT) of non-expert decision-makers (DTA). These theorems, however, are dependent on preconditions which appear extremely demanding to fulfil in real-world applications. This has led to considerable debate about the possibility to base arguments for democracy on them.[147] Moreover, the social choice

---

[142] Hélène Landemore, "Yes We Can (Make It Up on Volume): Answers to Critics", in: (2014) 26 *Critical Review* 1–2, 184–237, 187.

[143] Robert E. Goodin and Kai Spiekermann, *An Epistemic Theory of Democracy* (Oxford: Oxford University Press, 2018)

[144] Vermeule, "Collective Wisdom and Institutional Design", op. cit. 12, 346. Cass R. Sunstein, *A Constitution of Many Minds: Why the Founding Document Doesn't Mean What It Meant Before* (Princeton, N.J: Princeton University Press, 2009)

[145] Lu Hong and Scott E. Page, "Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers", in: (2004) 101 *Proceedings of the National Academy of Sciences* 46, 16385–16389, 16388; Scott E. Page, *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies* (Princeton, NJ: Princeton University Press, 2007), 163.

[146] Jean-Antoine-Nicolas de Caritat Marquis de Condorcet, *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* (Paris: De l'Imprimerie Royale, 1785)

[147] That voters are on average better than random (i.e., better than flipping a coin) on making decisions on any political questions is, for Estlund, implausible and renders the Theorem largely "irrelevant", cf. Estlund, *Democratic Authority*, op. cit. 27, 223–236; See also John Rawls, *A Theory of Justice* (Cambridge,

literature on judgment aggregation is not only exploring the 'wisdom of crowds' but also paradoxes and impossibility theorems emerging within group decision making.[148]

Critics of epistemic defenses of democracy argue that it is impossible to conclusively establish democracy's superiority based on *a priori* models.[149] Moreover, there are epistemic critiques of democracy which "*tend* to be a posteriori or empirical"[150]. However, Landemore convincingly argues that those epistemic critiques of democracy themselves endorse *a priori* models such as Down's theory of rational ignorance and Caplan's theory of rational irrationality.[151] Moreover, she argues that those critiques cannot disprove our expectations as to what certain decision-making systems may achieve, given that the preconditions of our models are met.[152]

While in the institutional dimension of PEA, voter ignorance may be negligible insofar as we are not trying to make the case for direct democracy, the criticism of the models being *a priori* requires further reflection. The position taken here is that the very *a priori* character of the DTA and the CJT provides us with a meaningful positive benchmark of institutional competence. The theorems can give us an idea of what drives group competence and what frustrates it. For example, shared cognitive distortions such as biases hinder the emergence of group wisdom. Thus, the empirical study of peoples' – including judges' – decision making is directly linked to our positive benchmark. Our benchmark, however, is not limited to merely negative institutional design approaches as suggested by Elster.[153] Potentially, aggregation holds the promise of cancelling out intellectual errors and cognitive biases of the decision-making group, at least as long as they are not all positively correlated.

MA: Harvard University Press, 1971), 358; Fabienne Peter, *Democratic Legitimacy* (London: Routledge, 2011); cf. also the Symposium on Landemore in Critical Review, op. cit. 141.

[148] See: Christian List, "The Theory of Judgment Aggregation: An Introductory Review", in: (2012) 187 *Synthese* 1, 187: 179–207; Lewis A. Kornhauser and Lawrence G. Sager, "Unpacking the Court", in: (1986) 96 *Yale Law Journal* 1, 82–117.

[149] Cf. Jason Brennan, "How Smart is Democracy? You Can't Answer that Question a Priori", in: (2014) 26 *Critical Review* 1–2, 33-58.

[150] Ibid., 35, my italics.

[151] Landemore, "Yes We Can (Make It Up on Volume)", op. cit. 142, 197; referring to the critiques by Brennan, "How Smart is Democracy?", op. cit., 149, 33–58 and Ilya Somin, "Why Political Ignorance Undermines the Wisdom of the Many", in: (2014) 26 *Critical Review* 1–2, 151–169.

[152] Similarly: Landemore, "Yes We Can (Make It Up on Volume)", op. cit. 142, 199.

[153] See Introduction.

Importantly, however, the purpose of this paper is not to discuss "the intrinsic persuasiveness" [154] of the mechanisms said to produce collective wisdom. Vermeule argues for an attitude of "qualified skepticism" towards those mechanisms because of the narrowness of the conditions under which they apply. Instead, he asks: "Under conditions in which collective wisdom can exist, how should institutions be structured to generate and exploit that wisdom?"[155]

In light of the above, PEA suggests a framework under which to anticipate this social epistemological research for the legitimacy of ICs. The epistemic theorems are not understood here as proof of the (moral) value of democratic rule. Instead, they can provide an institutional benchmark for judicial and other expertise-driven authority. Landemore does not claim democracy to be omnipotent in solving problems. She does not deny the necessity to delegate certain tasks to – democratically authorized – experts such as central bankers.[156] This is what the framework of PEA seeks to capture: public epistemic authorities are practical authorities whose authority is epistemically justified following a public decision-making process. Because we publicly empower these institutions due to their alleged epistemic superiority, we can also reverse-engineer their legitimacy by checking whether we have reasons to believe that they are actually better at solving the task they are meant to solve. ICs fall under this category. We install them to reliably determine correctly whether or not the law has been breached.

For scholars of international adjudication this brings in interesting new variables through which to study the legitimacy of ICs: For example, should the diversity of legal education on the bench of ICs amount to cognitive diversity in problem-solving models, then ICs could have an advantage over national courts, given that the latter are less cognitively diverse due to less diversity in their legal education. By the same logic, ICs could have to scale back their intensity of review towards national and supranational parliaments, as this paper continues to discuss.

As mentioned, those theorems are highly demanding in terms of their applicability. This paper proceeds without a thorough discussion of the conditions which have to be met for the DTA and the CJT to work. It only mentions them where the argumentation

---

[154] Vermeule, "Collective Wisdom and Institutional Design", op. cit. 12, 338.
[155] Ibid.
[156] Landemore, "Yes We Can (Make It Up on Volume)", op. cit. 142, 189.

requires it. Two broader conditions shall be mentioned here already: Technically, the combination of divergent opinions has to proceed by (majority) voting and not through consensus. Second, the members in the group have to share a common goal, a common objective function. Hence, we may want to bracket decisions under fundamental values disagreement.[157]

b. Institutional Implications of PEA

The framework of PEA carries institutional implications. It lends itself to making suggestions for institutional (re-)design and drafting principles for judicial reasoning. Institutional redesign can reach from rewriting the rules of evidence at ICs (of which there are strikingly few, as mentioned above) to changing the composition of the bench of ICs. For example, if the DTA holds, then there is a corresponding design principle which aims at increasing the cognitive diversity on the bench. One way to implement such a principle would be to open the international bench to non-lawyers, an approach which Vermeule favors for the U.S. Supreme Court.[158] In fact, as will be shown below, international adjudication is (and has been) open to include non-lawyers.

The institutional redesign of ICs, however, requires treaty change, the prospects of which are uncertain. An alternative way to suggest changes to the institutionalization of judicial authority is to propose principles which are meant to guide judges' reasoning along epistemic lines. Principles of epistemic deference and of judicial restraint are a key concern for institutional approaches to judicial authority. In the following section, we will thus discuss epistemic principles of judicial reasoning based on the very same mechanisms that ground our recommendations for institutional redesign.

---

[157] This is following Landemore, who brackets value diversity and assumes a general 'common good' orientation in voting, cf. ibid., 199. The formal theorems Landemore relies on require her to assume away diversity of fundamental values because the individuals in the decision-making group need to share a common goal or minimal background consensus. This, she writes, allows her to offer

> "a useful normative benchmark in some respects. Thus, if such an idealized democracy turns out to have desirable properties, we can ask ourselves what it would take to make real democracies more like it or, if that is impossible how we can retrieve some of the ideal properties through further 'countervailing deviations'"*

* ibid., 200; citing: Estlund (2014): "Fighting Fire with Fire (Departments)", Paper presented at the Princeton University Workshop on "Epistemic Dimensions of Democracy Revisited", April 30.

[158] Vermeule, "Collective Wisdom and Institutional Design, op. cit. 12.

III. Applying the Framework of Public Epistemic Authority

Let us complete our exploration of the institutional dimension of IC legitimacy by applying the framework of PEA to institutional (re-)design and judicial reasoning. Much of the following work will focus on the CJEU. Admittedly, the supranational European judiciary is somewhat of an outlier in the world of international adjudication due to the EU's high level of institutional integration. However, the existence of a European Parliament and a European technocratic executive (i.e., the European Commission) provide plenty of room for discussing the allocation of decision-making authority amongst international institutions. Moreover, the CJEU's jurisdiction overlaps with that of national courts as well as other international courts (i.e., the ECtHR). The CJEU thus fits well the research agenda of this paper.

1. Institutional (Re-)Design

This section discusses three issues concerning the institutional design of ICs: the cognitive diversity of the bench (a.), the size of the judicial chamber or panel (b.), and the appointment of experts and the appreciation of their expert opinions (c.). These are, of course, neither the only epistemic concerns in the design of an IC.[159] Nor would satisfying our epistemic demands fully legitimize an IC's public authority along all dimensions of legitimacy. These three issues, however, are a first institutional step to address the challenges to judicial authority.

Before we begin, one caveat is due: Below we are going to compare institutions based on mechanisms of collective wisdom. For the sake of the argument, we will assume that such comparisons are feasible. As regards statistical aggregation, however, technically speaking comparing two different groups with different sizes and different levels of

[159] A full account of institutional redesign would require considering the types of task which the Court most commonly engages in and then to select the organization and procedures which could allow its Judges to discharge these best. Cf. Chalmers, "Judicial Performance, Membership, and Design at the Court of Justice", in: Michal Bobek (ed.), *Selecting Europe's Judges: A Critical Review of the Appointments to the European Court* (Oxford: Oxford University Press, 2015), 51–77, 55.

individual competence may require more than the mere application of the CJT.[160] In line with our general careful approach towards the real-world applicability of mechanisms of collective wisdom, we will proceed *as if* such comparisons were feasible, in order to learn more about the direction in which institutional legitimacy flows.

a. Optimizing the Cognitive Diversity of the Bench

A design principle which seeks to optimize – instead of maximizing[161] – the cognitive diversity of the bench is the result of an application of the DTA. So long as no one in the decision-making group has the perfect answer to the problem at hand, the DTA tells us that the competence of the group to track the right answer increases when their cognitive models for finding solutions differ. The practical implications of optimizing the cognitive diversity of the bench lies in redesigning the judicial selection process. We will explore two variables for an optimized cognitive diversity: the national diversity of IC judges and its working language as well as the introduction of non-lawyers.

As mentioned, we will focus for now on the EU judiciary. Since the Treaty of Lisbon, the "Court of Justice of the European Union" (in here: CJEU) refers to the whole institution, including at the top-tier the "Court of Justice" (in here: CoJ), which has previously been known as the "European Court of Justice" (ECJ) and is still often referred to as such in academic and professional circles[162]. The middle tier is comprised of the "General Court" (in here: GC), formerly known as the "Court of First Instance" (CFI). The bottom-tier features "specialized courts", which were previously referred to as "specialized panels" and which once featured a now-abolished Civil Service Tribunal (CST). Hence, the current EU judiciary is a two-tier system. The CoJ consists of one

---

[160] Cf. Goodin and Spiekermann, *An Epistemic Theory of Democracy*, op. cit. 143, 20. See further, the Grofman Dummkopf-Witkopf Theorem, going back to: Bernard Grofman, "Judgmental Competence of Individuals and Groups in a Dichotomous Choice Situation: Is a Majority of Heads Better Than One?", in: (1978) 6 *Journal of Mathematical Sociology* 1, 47–60; see: Bernard Grofman, Guillermo Owen and Scott L. Feld, "Thirteen Theorems in Search of the Truth", in: (1983) 15 *Theory & Decision*, 261–278, 265.

[161] Cf. Vermeule, "Collective Wisdom and Institutional Design", op. cit. 12.

[162] Cf. Marie-Pierre Granger and Emmanuel Guinchard, "Introduction: The Dos and Don'ts of Judicial Reform in the European Union", in: Emmanuel Guinchard and Marie-Pierre Granger (eds.), *The New EU Judiciary: An Analysis of Current Judicial Reforms*" (Alphen aan den Rijn: Kluwers Law International, 2018), 1–18, 1.

Judge[163] from each Member State who are assisted by Advocates General.[164] The GC is comprised of two Judges per Member State.[165] The CoJ's Judges and Advocates General are chosen from

> "persons whose independence is beyond doubt and who possess the qualifications required for appointment to the highest judicial offices in their respective countries or who are jurisconsults of recognised competence".[166]

For the Judges of the GC, their appointment requires likewise independence "beyond doubt", but only demands "the ability required for appointment to high judicial office".[167]

Since the Treaty of Lisbon, the selection of Judges and Advocates General has changed. Before Lisbon, the respective national governments selected their candidates internally, which were then appointed by common accord.[168] The national selection criteria were fairly unclear, often non-transparent, and not subject to open applications; the common accord was more formal procedure than a substantive testing of the candidates.[169]

---

[163] For the remainder of this paper, the capitalized "Judge" refers to a judge of the Court of Justice of the European Union. Likewise, the capitalized "Advocate General" refers to this position at the same court.

[164] As far as Article 19 (2) 1 Treaty on European Union (TEU) goes, the CoJ consists of one Judge from each Member State and is assisted by Advocates General. The GC "shall include at least one judge per Member State". Article 19 (2) 3 TEU claims that the Judges of the CoJ and the GC as well as the Advocates General are chosen from "persons whose independence is beyond doubt" and who satisfy further conditions set out in the TFEU. The Judges and Advocates General are then appointed by "common accord" of the Member States' governments for a renewable term of six years. The final selection decision is made by the representatives of the governments of the Member States, cf. European Council, Council Decision 2010/124/EU relating to the operating rules of the panel provided for in Article 255 of the Treaty on the Functioning of the European Union, February 25, 2010, Annex, para 8. Article 252 (1) Treaty on the Functioning of the European Union (TFEU) then sets the number of Advocates General at the CoJ to eight, though subject to potential increase through the Council. In fact, the CoJ currently is comprised of 28 Judges and eleven Advocates General, cf. http://curia.europa.eu/jcms/jcms/Jo2_7024/en/ (last accessed on: November 30, 2018).

[165] According to Art 254 (1) 1 TFEU, the number of the Judges at the GC is determined by the Statute of the Court of Justice of the European Union – which currently sets it at two per Member State, cf. https://curia.europa.eu/jcms/jcms/Jo2_7033/en/ (last accessed on: November 30, 2018).

[166] Art 253 (1) TFEU.

[167] Art 254 (2) 1 TFEU.

[168] Michal Bobek, "The Court of Justice of the European Union", in: Damian Chalmers and Anthony Arnull, *The Oxford Handbook of European Union Law* (Oxford: Oxford University Press, 2015), 153–177, 163.

[169] Ibid., 164; See further: Anthony Arnull, *The European Union and its Court of Justice*, 2nd edition, (Oxford: Oxford University Press, 2006), 23–24.

With Lisbon, a new '255 Panel' as been created,[170] its name following Article 255 Treaty on the Functioning of the European Union (TFEU). Comprised of jurists, representative of the different legal systems and different European regions,[171] the 255 Panel provides opinions on the individual candidates for CoJ and GC Judges and Advocates General proposed by a Member State to the representatives of the governments of all Member States.[172] While the Panel cannot make its own nominations, it nevertheless intervenes before the governments decide together.[173] Before the Lisbon Treaty took effect, the appointment mechanism was designed and practiced exclusively on mutual trust between the Member States. This led to the norm of not voicing negative views on a nominee.[174]

According to the latest Activity Report of the Panel, since its creation in 2010 it has delivered a total of 147 opinions.[175] Of the 147 candidates assessed, 61 were for the office of Judge or Advocate General of the CoJ and 86 for the office of Judge of the GC. Of these candidates, 74 were proposed for renewal of their term of office at the CoJ (39) or the GC (35). 73 candidates for a first term of office were also assessed, including 22 for the CoJ and 51 for the GC. 14 of the 147 opinions delivered since 2010 have been unfavorable. No unfavorable opinions have been delivered on candidates for the renewal of a term of office. This means that 19.2% (14 out of 73) of the opinions on candidates for a first term of office were unfavorable.[176] The Panel's opinions are non-binding, but so far have been

---

[170] Bobek, "The Court of Justice of the European Union", op. cit. 168 164.

[171] The composition of the Panel from 2014 to 2018 has been: Mr. Luigi Berlinguer, first Vice-Chair of the European Parliament's Committee on Legal Affairs; Ms. Pauline Koskelo, Judge of the European Court of Human Rights and former President of the Supreme Court of Finland; Lord Mance, Judge and, since 2017, Deputy President of the Supreme Court of the United Kingdom; Mr. Jean-Marc Sauvé, Vice-President of the Council of State of France; Mr. Christiaan Timmermans, former President of Chamber of the Court of Justice of the European Union; Mr. Andreas Voßkuhle, President of the Federal Constitutional Court of Germany; and Mr. Mirosław Wyrzykowski, former Judge of the Constitutional Court of Poland; cf. *Fifth Activity Report* (Article 255 Panel), 4.

[172] Cf. Council Decision 2010/124/EU, relating to the operating rules of the panel provided for in Article 255 TFEU, OJ L 50.

[173] Bobek, "The Court of Justice of the European Union", op. cit. 168, 164.

[174] Jean-Marc Sauvé, "Selecting the European Union's Judges: The Practice of the Article 255 Panel", in: Bobek, *Selecting Europe's Judges*, op. cit. 159, 78–85, 79.

[175] CJEU, *Fifth Activity Report of the panel provided for by Article 255 of the Treaty on the Functioning of the European Union* (Article 255 Panel), published on February 28, 2018, available under: https://curia.europa.eu/jcms/upload/docs/application/pdf/2018-05/5eme_rapport_dactivite_du_c255_-_en_final_-_public.pdf (last accessed on: October 6, 2018). Technically, the panel is not continuous but reconstitutes itself every four years, cf. ibid.

[176] Ibid., 15.

followed by the governments.[177] In the literature, the Panel has been regarded as effective in reducing arbitrariness in judicial appointments to the CJEU.[178] However, it should be noted that the more general research on judicial performance evaluation is more critical of judicial panels.[179]

The Panel arguably sets higher standards for office than the requirements in the Treaties. It appears to apply the following six sets of criteria:

> "the candidates' legal capabilities; their professional experience; their ability to perform the duties of a Judge; their language skills; their ability to work as part of a team in an international environment in which several legal systems are represented; whether their independence, impartiality, probity and integrity are beyond doubt; in future, the panel will also take into account the physical capacity of candidates to carry out demanding duties which require considerable personal investment."[180]

The Panel thus not only assesses the candidates' independence, but also their competence and expertise.[181] Interestingly, the Panel does not assess the "scope and comprehensiveness of candidates' legal expertise, particularly with regard to European Union law."[182] The failure to answer a precise question on EU law does not automatically lead to a negative view on a candidate's ability. While valuing "diverse opinions", the Panel expects nothing more than an adequate basic knowledge of EU law and the ability to analyze and reflect on its general issues, a requirement "met without difficulty by a high-level generalist who is not specialised in Union law."[183] However, insufficient

---

[177] Bobek, "The Court of Justice of the European Union", op. cit. 168, 164. Whether the compliance is due to the expertise of the members of the Panel or the requirement of unanimity in the common accord is an open question, cf. ibid.

[178] See, for example: Tomas Dumbrovsky, Bilyana Petkova, and Marijn Van der Sluis, "Judicial Appointments: The Article 255 TFEU Advisory Panel and Selection Procedures in the Member States", in: (2014) 51 *Common Market Law Review* 2, 455–482.

[179] Cf. Jennifer K. Elek and David B. Rottman, "Improving Judicial-Performance Evaluation: Countering Bias and Exploring New Methods", in: (2013) 49 *Court Review* 3, 140–144; Nuno Garoupa and Tom Ginsburg, "The Comparative Law and Economics of Judicial Councils", in: (2009) 27 *Berkeley Journal of International Law* 1, 52–83; Lawrence Baum, *Judges and Their Audiences: A Perspective on Judicial Behavior* (Princeton, NJ: Princeton University Press, 2006), Chapter 6.

[180] Cf. *Fifth Activity Report*, op. cit. 175, 25.

[181] Sauvé, "Selecting the European Union's Judges", op. cit. 174, 79.

[182] *Fifth Activity Report*, op. cit. 175, 29.

[183] Ibid.

knowledge of EU law and inadequate legal capabilities have in fact led to unfavorable opinions by the Panel.[184] Unfavorable opinions have been delivered for instance where the candidates' length of high-level professional experience was found "manifestly too short" and this fact could not be compensated for by "exceptional or extraordinary legal capabilities".[185] The panel has also occasionally noted the complete absence of any professional experience relevant to EU law.[186]

Some argue that the role of experts in appointing judges should be narrowly tailored.[187] Especially deliberative democratic accounts of judicial authority must be worried about expert discourse being unnecessarily opaque and unfit for public discussion. As Armin von Bogdandy and Christoph Krenn argue, in terms of what makes a good Judge, the Panel "cannot help but make important value judgments, which it needs to hide under the guise of expertise"[188]. For PEA, the role of experts in appointing the Judges is merely instrumental insofar as it might or might not help to raise the epistemic reliability of the Court as a whole. At this point, the difference between an individualist paradigm of judicial competence as instituted in the EU's judicial selection system and a collectivist paradigm such as that of PEA becomes decisive. What matters significantly under the latter is the question of composition, which the Panel explicitly excludes to be dealing with.[189] However, if it should prove to be true that under the Panel domestic judges have a better chance of being successfully selected than academics or diplomats as their profiles are more comparable to those of the Panel members,[190] this could likely

[184] Ibid., 30.
[185] Ibid., 29.
[186] Ibid.
[187] See, for example, from a German perspective: Andreas Voßkuhle and Gernot Sydow, "Die demokratische Legitimation des Richters", in: (2002) 57 *JuristenZeitung* 14, 673– 682.
[188] Armin von Bogdandy and Christoph Krenn, "On the Democratic Legitimacy of Europe's Judges: A Principled and Comparative Reconstruction of the Selection Procedures", in: Bobek, *Selecting Europe's Judges*, op. cit. 159, 162–180, 175.
[189] CJEU, *Third Activity Report of the Panel Provided for by Article 255 of the Treaty on the Functioning of the European Union* (Article 255 Panel), published on December 13, 2013, available under: https://curia.europa.eu/jcms/upload/docs/application/pdf/2014-02/rapport-c-255-en.pdf (last accessed on: December 10, 2018), 11.
[190] Suggested in: Dumbrovsky et. al., "Judicial Appointments", op. cit. 178. In this regard, Michal Bobek's take on the Court's selection process is further illuminating. He writes:

> "*Functionally*, the abstract law making carried out by the ECJ [...] requires analytical minds that are able to rise above a single case and a national judicial routine, and able to see the bigger European picture. *Culturally* and historically, insisting on a higher court being composed of (mainly or wholly) senior judges who are believed to be the only

negatively influence the cognitive diversity of the bench. We thus need to take a closer look at the judicial selection process. We can utilize approaches in the field of behavioral law and economics to sophisticate our analysis. As mentioned, the complete research paradigm of the PEA framework combines our *a priori*, aggregative benchmark with the input of *a posteriori* behavioral variables. While the following paragraphs do not present a full exercise of the Behavioral Test of Legitimacy as envisioned by PEA, they illuminate further what kind of behavioral input is informative in the institutional dimension of legitimacy.

Angela Huyue Zhang's has presented one of the first behavioral studies of the CJEU personnel. Amongst other issues, her study discusses judicial selection as a performance-sensitive aspect.[191] Zhang highlights the fact that there is no common market for Judges: the Member States bring forward their own candidates.[192] Since Member States never disagree with each other's nomination, each Member State practically appoints its own Judges[193] – following its own, possibly opaque and political judicial appointment process.[194]

> competent persons to understand the business of judging is the reflection of one particular legal tradition within the Union. Conversely, other legal traditions within the Union are much more open towards higher courts also being composed of academics."*

* Bobek, "The Court of Justice of the European Union", op. cit. 168, 166-167, orig. ital.

[191] The other aspects are compensation and tenure, cf. Angela Huyue Zhang, "The Faceless Court", in: (2016) 38 *University of Pennsylvania Journal of International Law* 1, 71–135, 81–93. Overall, within her framework she lists six factors as negatively influencing the Court's performance: Zhang lists six factors adverse to performance: First, there is the risk of political appointees, attracted by the high judicial salaries at the Court and a lack of procedural safeguards. "As a consequence, some judges who are selected are not competent to perform and are dominated by their référendaires", ibid., 71. Second, the dependence on the *référendaires* is further increased by the high turnover of judges. Third, the *référendaires*, in turn, are drawn from a relatively closed social network, displaying an inefficient labor market which prolongs their stay at the Court. Fourth, a "revolving door between the Court and the European Commission raises serious conflict issues, as the Commission is able to [...] gain a comparative advantage in litigation", ibid., 72. Fifth, the Court issues single, collegial decisions which encourages free-riding and suppresses dissent. Sixth, "the division of labor between the General Court and the Court of Justice could lead to divergent incentives for judges working at different levels of the Court", ibid., 72.

[192] Ibid., 81; citing: Sally J. Kenney, "Breaking the Silence: Gender Mainstreaming and the Composition of the European Court of Justice", in: (2002) 10 *Feminist Legal Studies* 3–4, 257–270.

[193] Zhang, "The Faceless Court", op. cit., 191, 81; citing: Kenney, "Breaking the Silence", op. cit. 192.

[194] Ibid.; referencing: Henri de Waele, "Not Quite the Bed that Procrustes Built: Dissecting the System for Selecting Judges at the Court of Justice of the European Union", in: Bobek (ed.), *Selecting Europe's Judges*, op. cit. 159, 24–51.

According to Zhang's study of 2016, over 65% of the EU Judges have worked in government before they joined the CJEU (67% at the CoJ and 57% at the GC).[195] For Zhang,

> "[t]he preference for government officials is not surprising. As the secret deliberations rule prevents nominating states from monitoring the voting preference of their appointees, appointing governments are more prudent in choosing the candidates that they believe will act in their interest."[196]

Without any public hearing during the appointment process, the only source of public information on the Judges is their profiles presented by the CJEU. There is no mandatory disclosure rule and thus many of the profiles are incomplete: in Zhang's study, almost 77% of the profiles contain missing information about the Judge's education background (26% of profiles provide no education information at all).[197] Another 16% of the profiles lack proper information about the Judge's primary work experience and of roughly 18% of the Judges it is unknown what last position they held before joining the EU's bench.[198] Zhang herself admits the difficulty of inferring the Judges' qualification for the job from the sparse information given by the CJEU. However, as she sums up her view on the loose criteria established by Articles 253 and 254 TFEU, the three "important skills" needed to function effectively and efficiently at the CJEU are: knowledge of EU law, legal and research skills, and the command of the French language.[199]

Now, the question is whether the CJEU's epistemic reliability would be higher if a common market for Judges (and Advocates General) would be established and the one Judge per Member State rule would be abolished. At first sight, these three 'important' skills could be maximized if the CJEU could simply draw from the best applicants within the EU. The benefits of cognitive diversity, however, suggest reconsidering this move. Assuming that the applicability conditions of the DTA are met, if we add an additional decision maker to a group, what matters most is that she brings in a different perspective,

[195] Ibid.
[196] Ibid., 82.
[197] Ibid., 82-83. At the time of her study, five out of seven appointees from Denmark provided no educational information whatsoever. Most judges from Denmark come from government positions, ibid.
[198] Ibid.
[199] Ibid., 83.

not necessarily that she is a better decision maker. An individualist paradigm of legal expertise cannot express such subtleties.

Even if the applicability conditions of the DTA are met, we have to consider the fact that the EU Judges usually sit in chambers. Three (or five) Judges are an extremely small group in which the three (or five) best performers are likely to outperform the three (or five) most diverse ones. Yet if the three or five best are very similar in their skill set, diversity may count again. Given that the selection criteria at the CJEU all screen for generalist legal skills (and the command of French), the diversity within a chamber must be looked at from two angles. As regards their legal skills, a chamber of three Judges from different legal traditions appears more diverse than a chamber filled with three French judges. As regards their ability to draw on non-legal skills and extra-legal knowledge, even a chamber which is diverse in legal skills appears to be little diverse in skill overall. Whether chambers dealing with empirical epistemic uncertainty are thus likely to benefit from adding an economist or a scientist will be further discussed below.

What we can conclude for now is that the system of one Judge per Member State introduces some diversity of legal skill to the CJEU. From a collectivist paradigm of judicial competence, this is a welcome effect which cannot be accounted for by the individualist paradigm reflected in the positive EU law governing the Union's judicial appointments. Of course, if the motivation behind the one Judge per Member State-rule is supposed to make sure that domestic preferences are represented on the bench, PEA is blind towards this matter, as its model of adjudication is one of judgment aggregation.[200]

There is, however, another difficulty with an institutional design approach which sacrifices (some) expertise for (some) diversity in appointing EU Judges. According to Zhang's study, based on expert interviews, the less capable an EU Judge is, the more he or she is dependent on his or her *référendaires* – the Court's legal secretaries – to carry

[200] We may see ourselves supported by a study by Damian Chalmers, who divides the history of the CoJ into four different time periods, each dominated by a particular task the court gave itself with no explicit mandate. Chalmers then compares these tasks in time with the professional background of each Judge and Advocate General working at the CoJ during these periods. Notwithstanding the methodological difficulties of the assumptions at play here, Chalmers draws the conclusion that the iconic (activist) judgments of the CoJ came in time periods where the Court was dominated by academics and judges and less by civil servants and practitioners. Cf. Chalmers, "Judicial Performance, Membership, and Design at the Court of Justice", op. cit. 159.

out the judicial functions of the office.[201] "As a consequence, the voices of *référendaires* are amplified, and in some instances they effectively become the Judges behind the scenes."[202] At each level of the CJEU, every Judge or Advocate General is entitled to three *référendaires*; a fourth *référendaire* is granted to some Judges who assume management responsibilities. At the CoJ, the Judges are further entitled to the additional help from *administrateurs juristes* – lawyers who do not work on cases directly.[203] Zhang obtained data provided by the Court, according to which in March 2015, there were 123 *référendaires* and 22 *administrateurs juristes* working at the CoJ and 94 *réféfendaires* at the GC.[204]

Michal Bobek – Advocate General at the CoJ since 2015 – indirectly gives further weight to Zhang's findings. He writes:

> "They [the *référendaires*] are not members of the ECJ or the GC; they are *just* assisting individual judges or Advocates General. What precisely 'assisting' means will depend on the individual court member and the working habits within the respective judicial chambers. In practical terms, 'assisting' may mean anything between researching the case law and writing memoranda and acting as a fully-fledged ghost writer, drafting but never signing a judgment or Opinion. Whatever the case may be, it is clear that writing a judicial decision, especially at the supreme level, is nowadays a collective enterprise, both across the individual chambers as well as within."[205]

Only recently, the *référendaires* have been studied at some depth (or rather, at all).[206] Zhang's study presents a novel way of studying their background: In 2015, Zhang used the career platform LinkedIn to hand-collect the background data for 74 current *référendaires*, 31 at the CoJ, 43 at the GC. This amounts to 25% of current *référendaires* at the CoJ and 46% at the GC. Additionally, Zhang used LinkedIn to hand-collect the

---

[201] Zhang, "The Faceless Court", op. cit. 191, 85.
[202] Ibid.
[203] Ibid., 77.
[204] Ibid.
[205] Bobek, "The Court of Justice of the European Union", op. cit. 168, 168, orig. ital.
[206] Cf. Stephane Gervasoni, "Des référendaires et de la magistrature communautaire", in: Jean-Pierre Puissochet, *L'État souverain dans le monde d'aujourd'hui* (Paris: Editions A. Pedone, 2008), 105–126.

background data of 103 former *référendaires*.[207] There are several methodological problems with this approach, all of which cannot be discussed here. But since her only source is LinkedIn, it is likely that some groups are underrepresented. Former *référendaires*, who are now working in public institutions are less incentivized to use LinkedIn than those who seek employment in private practice, due to each sector's recruitment practices. Her interviewees indicate that "a sizeable portion of référendaires are administrative judges from France but none of them appear in the samples"[208]. Moreover, the use of LinkedIn as a career platform may vary from Member State to Member State, making it for some groups of *référendaires* more likely to have an account with this platform than others. Notwithstanding these difficulties, one finding of her study is relevant for our discussion. There is a significant dominance of French legal culture amongst the *référendaires* at the CJEU. Over 42% of *référendaires* at the CoJ are from France, Belgium and Luxembourg. At the GC their number adds up to 49%. When it comes to legal education and not merely country of origin as a crude proxy for legal tradition, the numbers are even stronger: 79% at the CoJ and 83% at the GC are educated in law schools in France, Belgium, and Luxembourg. After all, the *référendaires* have little common law training.[209] It does not require a bold inferential step to assume that the French dominance correlates with the working language of the Court being French.

We must thus concede that not every Judge will introduce more or, rather, an equal amount of diversity in skill *qua* legal culture if she is dominated by her *référendaires* which in turn may be predominantly trained in the French legal culture.[210] Under the framework of PEA, changing the working language of the EU courts from French to English thus suggests itself.

As mentioned above, another way to increase the cognitive diversity of the bench is to add non-lawyers to the bench. The idea is not exclusively owned by epistemic theorists of

---

[207] Zhang, "The Faceless Court", op. cit. 191, 96.
[208] Ibid., 95.
[209] Ibid., 96-98.
[210] Compensation plays a role here, too. Zhang's study notes that the Judges' salary is "currently set at a level that far exceeds the pre-existing salary for the vast majority of national Supreme Court judges", ibid., 91. While the resulting pay rise through the appointment to the Court might attract more qualified candidates, "but also those less genuinely interested in judging than in the perks and benefits the job brings.", ibid. In the study, several interviewees stated to have observed that some judges who received significant pay increases were political appointees, incompetent to perform their duties and dominated by their *référendaires*, cf. ibid.

public authority. At times, deliberative accounts have made similar suggestions. Jan Komárek claims that some theorists suggest that concentrated constitutional courts[211] are to be preferred over having constitutional adjudication dispersed among the whole judiciary. Summarizing their arguments, Komárek lists greater professional diversity as one of the reasons: they include "members with different backgrounds, not just lawyers or even career judges, as is still usual in continental Europe."[212]

Under PEA, the introducing non-lawyers is of course discussed from the point of view of epistemic reliability. The assumption, circulated by Vermeule, is that the epistemic gains of including more than zero non-lawyers outweigh the epistemic losses.[213] On the one side, if a scientist with the relevant knowledge in the field was added to cases under scientific uncertainty discussed above, her addition would improve the Court's collective competence in those cases and likely reduce the correlation of bias. The latter would even be had if the scientist does not have the relevant expertise.[214] On the other side, the collective legal competence of the Court would decrease and there may be further costs involved with including a non-lawyer to the bench which are not captured by the rationale of the DTA (or the CJT, see below). There is further uncertainty towards the magnitude of both costs and benefits. In such situations, Vermeule is right when he writes about the U.S. Supreme Court that:

> "Under uncertainty, I suggest, it is a poor strategy of institutional design to adopt an extreme solution, filling all [...] slots with lawyers. Rather, on marginalist grounds, the sagest guess is that the epistemic gains of adding the first nonlawyer will exceed the epistemic costs of subtracting the ninth lawyer."[215]

A stronger version, which would in practice be much more difficult to obtain, is a requirement of dual competence for lawyers. Scott Brewer's two-hat solution is such an

---

[211] Most Member States have so-called concentrated constitutional courts, i.e. "specialised judicial institutions empowered to review the constitutionality of the exercise of public power", cf. Jan Komárek, "The Place of Constitutional Courts in the EU", in: (2013) 9 *European Constitutional Law Review* 3, 420–450.
[212] Ibid., 425.
[213] Vermeule, "Collective Wisdom and Institutional Design", op. cit. 12, 357.
[214] Cf. ibid.
[215] Ibid., 357-358.

attempt, which we showed above to be unnecessarily rooted in an individualistic paradigm of judicial competence.

According to the current laws governing the judicial selection of the CJEU, it is almost certain that only lawyers will ascend to its bench. However, it is not uncommon for European high courts to feature non-lawyers. The French *Conseil Constitutionnel* does not require its judges to have any particular professional qualification.[216]

Historically, not even the first Judges of the CJEU were all lawyers.[217] The current CJEU would most likely be skeptical towards any suggestion of breaking the monopoly of lawyers. As regards its case law featuring complex economic assessments, Marc Jaeger, President of the GC, acknowledges the epistemic benefits of requiring Judges' to have economic training, but holds any requirements of specialized, non-legal knowledge to go against the "spirit of the treaty":

> "While the nomination of judges with a specific understanding of economics would undisputedly not do any harm, the first and foremost abilities required from a judge is his or her independence and capacity to discern what is relevant and to disregard what is not, whatever the complexity or the matter of the question submitted to his or her decision. […] To require that judges shall dispose of a specific expertise in a given field – economics or something else (e.g. intellectual property, computing science, chemicals) – would thus risk going beyond the letter and the spirit of the treaty. Additionally, from a practical and political point of view, one can hardly see how the margin of discretion enjoyed by Member States when proposing their candidates could be restrained in such a way."[218]

[216] The U.S. Supreme Court has not always been almost exclusively recruited from the bench of federal courts, cf. Terri L. Peretti, "Where Have All the Politicians Gone - Recruiting for the Modern Supreme Court", in: (2007) 91 *Judicature* 3, 112–122. Judges can be poor aggregators of political preferences where constitutional law requires preference aggregation, cf. Vermeule, "Collective Wisdom and Institutional Design", op. cit. 12.

[217] Ditlev Tamm "The History of the Court of Justice of the European Union Since its Origin", in: Court of Justice of the European Union, *The Court of Justice and the Construction of Europe: Analyses and Perspectives on Sixty Years of Case-Law* (The Hague: T.M.C. Asser Press, 2013), 9–35, 18, fn. 30.

[218] Marc Jaeger, "The Standard of Review in Competition Cases Involving Complex Economic Assessments: Towards the Marginalisation of the Marginal Review?", in: (2011) 2 *Journal of European Competition Law & Practice* 4, 295–314, 311.

Jaeger's account gives us an impression of how unlikely the implementation of PEA-based epistemic recommendations of institutional design might be. However, the point of PEA is to provide a comparative test for institutional legitimacy. Its value for IC scholarship is not dependent on its institutional implementation. Moreover, not every suggestion of institutional design under PEA would require Treaty change, neither of the European Treaties as discussed here nor the treaties establishing other ICs.

Jaeger continues to argue that while a specialized tribunal could easily absorb specialized judges, the "generalist nature" of the GC would complicate matters. The expertise of the candidates for judicial office at the GC "must continue to ensure that the diversity of the backgrounds of judges mirrors the growing diversity of cases."[219] Admittedly, if all judges were economist-lawyers then it may be hard to see how that might improve the CJEU's fundamental rights adjudication. However, it is hard to see how introducing *one* economist (or economist-lawyer) would hurt the diversity of the bench rather than improving it. Jaeger further asks whether the judges should "not only be economists but also technicians, engineers, biochemists, plumbers?"[220] In fact, EU legislation is peppered with technical vocabulary and concepts, not only economic ones.[221] However, while the natural and social sciences are not alike, having one judge with extensive training in statistics would likely benefit the CJEU's understanding of both. Second, the point of cognitive diversity is to add people with different problem-solving models to the group. Admittedly, a plumber would most likely have too little to add the bench in a case like *Google Spain* or *Afton Chemical*. But can we say the same about an engineer, a technician, or a biochemist? At least for fact-sensitive values like the precautionary principle or the fact-sensitive balancing of privacy with access to information and economic interests, their diverse input is more likely to be epistemically beneficial than harmful.

Again, PEA is not only a successful framework if a social or natural scientist ascends to the bench of an IC. This misses its point. Let us therefore look at a design suggestion which operates below the threshold of Treaty change. We could, for example, recommend that every Judge recruit one scientifically trained *référendaire* to work in her cabinet.

[219] Ibid.
[220] Ibid., 312.
[221] Ibid.

Then the Judge whose *référendaire* has her expertise in (or closest to) the scientific field required for the case at hand would automatically become the Judge Rapporteur of that case. Our analysis above showed the tremendous influence a Judge's cabinet can have on her decision making. It is beyond the scope of this paper to fully flesh out an epistemic system for the CJEU which includes the *référendaires*. Suffice it to note, however, that the Court already has a system in which cognition is "distributed"[222]: between the Advocates General, the Judges, the *référendaires*, the parties and their submissions, as well as observations by the Member States and the Commission.

b. Judicial Chambers and Statistical Aggregation

Within a collectivist paradigm of judicial competence, the chamber (or panel) size of an IC quite obviously matters. At the CoJ, only one case has been assigned to the Full Court in the past five years.[223] For certain types of cases a Grand Chamber is formed, consisting of 15 Judges and presided over by the President of the CoJ and including the Vice-President of the CoJ and three of the Presidents of the chambers of five Judges.[224] The CoJ comes together in the formation of the Grand Chamber when a Member State or an institution of the Union that is party to the proceedings so requests.[225] It is the formation

[222] James Dillon recently introduced a collectivist epistemological paradigm for courts, Cf. Dillon, "Expertise on Trial", op. cit. 16. His suggestion follows accounts of so-called "distributed cognition", in which the cognitive task at hand "is broken down into components, which are given to different members of the group. Membership of the group […] is a matter of having a particular function within the overall system", Alexander Bird, "When Is There a Group That Knows? Distributed Cognition, Scientific Knowledge, and the Social Epistemic Subject", in: Jennifer Lackey (ed.), *Essays in Collective Epistemology* (Oxford: Oxford University Press, 2014), 42–63, 45.

Dillon's suggestion lies in introducing a scientific adjunct to the Court who gets assigned to the court in order to incorporate scientific expertise into courts' epistemic systems, cf. Dillon, "Expertise on Trial", op. cit. 16, 295-296. Dillon focuses on criminal law trials and to some extend civil procedures. His scientific adjuncts are employed accordingly: at the pretrial stage, the scientific adjuncts would decide the issues of expert qualification and gatekeeping. These decisions are only subject to review for "clear error" by the trial judge. The parties would have no veto over the expert review, cf. ibid., 296-297. Of course, 'science' is the same in constitutional law and in criminal law, but constitutional law cases are marked by normative contestation which may motivate us to proceed with some care in our suggestions of institutional design. The benchmark of epistemic reliability affords a thorough argument to be defensible.

[223] Cf. CJEU, *2017 Annual Report: Judicial Activity* (Luxembourg, 2018), 116. The Court sits as a Full Court where cases are brought in pursuant to Articles 228 (2), 245 (2), 247 or 286 (6) TFEU, cf. Article 16 (4) of the Statute of the CJEU. After hearing the Advocate General, the Court can decide to refer a case to the Full Court if it considers that the case is of "exceptional importance", cf. Article 16 (5) Statute of the Court of Justice of the European Union.

[224] Article 16 (2) Statute of the Court of Justice of the European Union.

[225] Article 16 (3) Statute of the Court of Justice of the European Union.

which is entrusted to assume the coherence of the case law and also to further develop the law.[226] The Grand Chamber was originally composed of only 13 Judges who were mostly presidents of chambers and now includes elected members, a fact which Bobek comments on as suggesting "that the rationale has moved from a more 'unity in case law' driven composition of the Grand Chamber to perhaps more 'representative' and 'democratic' considerations."[227] Of the cases in 2017 which have been assigned already, roughly 22% were assigned to a Grand Chamber.[228]

The most frequent assignment at the CoJ is that to a chamber of five judges (in 2017, roughly 56% of assigned cases went to a five-Judge chamber).[229] The chamber of five Judges appears to be the default mode, unless the case is believed to be a straightforward application of existing case law, then a small chamber of three judges is deemed sufficient[230] (roughly 21% of assigned cases in 2017)[231]. These new structures can breed new types of problems: a five-judge chamber may head off unchecked and create its own case law in a given area or inconsistencies might emerge across chambers' decisions, requiring intervention by the Grand Chamber.[232]

At the GC, there are chambers consisting of three or five Judges. The default mode is a chamber of three Judges. According to the 2017 data, almost 88% of cases were assigned to a chamber of three and not even 7% to a chamber of five.[233] As mentioned, the institutional design of the GC has been dramatically altered through the so-called 'Skouris reforms'. In December 2015, the EU legislature adopted a regulation to reform the judicial structure of the CJEU.[234] The reform aims at enhancing the efficiency of the EU's judicial system through a three-step increase of the number of judges at the GC as well as through

---

[226] Cf. Article 27 (1) of the Rules of Procedure; Bobek, "The Court of Justice of the European Union", op. cit. 168, 157. The Grand Chamber includes the president of the Court, the vice-president, and three presidents of chambers of five judges, cf. ibid.

[227] Ibid., fn. 19.

[228] Own calculation, based on: CJEU, *2017 Annual Report: Judicial Activity*, 116. The proportions of chamber allocation appear fairly constant over the past five years, cf. ibid.

[229] Ibid. A case is assigned to a chamber, "so far as the difficulty or importance of the case [...] are not such as to require that it should be assigned to the Grand Chamber", cf. Article 60 (1) Rules of Procedure.

[230] Bobek, "The Court of Justice of the European Union", op. cit. 168, 156-157. Cf. Article 16 (1) Statute of the Court of Justice of the European Union.

[231] Own calculation, based on: CJEU, *2017 Annual Report: Judicial Activity*, 116.

[232] Bobek, "The Court of Justice of the European Union", op. cit. 168, 157.

[233] CJEU, *2017 Annual Report: Judicial Activity*, 219.

[234] Regulation 2015/2422/EU of the European Parliament and of the Council of 16 December 2015 amending Protocol No 3 on the Statute of the Court of Justice of the European Union, OJ L 341.

a change in the Rules of Procedure of the GC.[235] Within the CJEU itself, the efficiency-based approach has been received with skepticism. As Advocate General Sharpston criticizes the changes: "in a difficult case, speed may come at the expense of quality"[236].

In fact, the CJEU has witnessed numerous redesigns and the recent Skouris reforms of the GC were accompanied by much debate, not only among observers but also within the EU judiciary.[237] First of all, there was significant disagreement over the numbers meant to prove the backlog problem, especially with regards to the GC. Nevertheless, two options of institutional design have been considered: creating a new specialized court for intellectual property cases or increase the number of GC Judges.[238] Then President of the CoJ, Vassilios Skouris opted for the second and at a public hearing of the European Parliament Committee of Legal Affairs in 2013 he discussed the reform with the rapporteur from the European Parliament. As Franklin Dehousse, then Judge at the GC and highly critical of the reform,[239] recounts the hearing:

> "The debate's main theme was the criteria upon which the additional judges would be appointed. According to the Court's President, too much emphasis had been placed on the criterion of nationality. Three key points, in order of importance, had to be considered. First of all, the criterion of the

---

[235] CJEU, *2015 Annual Report: Judicial Activity* (Luxembourg, 2016), 104. The reform currently proceeds as follows: "The General Court shall consist of: (a) 40 Judges as from 25 December 2015; (b) 47 Judges as from 1 September 2016; (c) two Judges per Member State as from 1 September 2019", cf. Article 48 of the Statute of the Court of Justice.

[236] Eleanor Sharpston, "Transparency and Clear Legal Language in the European Union: Ambiguous Legislative Texts, Laconic Pronouncements and the Credibility of the Judicial System", in: (2010) 12 *Cambridge Yearbook of European Legal Studies*, 409–423, 418.

[237] As Marie-Pierre Granger and Emmanuel Guinchard write, the reform

> "triggered unprecedented anger, frustration, and anxiety, not only among observers and users of the EU judiciary, but also within the CJEU's own ranks. Among other problems, the reform led to an acrid confrontation between the two main EU courts, the CJ and the GC. […] Many observers – academics or journalists – have been very critical of the substantive outcomes of the reform. Moreover, even those who take a more positive stance on the achievements admit that the manner in which the reform was handled irremediably tarnished the reputation and legitimacy of the Court."*

* Granger and Guinchard, "The Dos and Don'ts of Judicial Reform in the European Union", op. cit. 162, 2.

[238] Cf. ibid., 1–14.

[239] Dehousse has commented the reforms in a total of three papers, proposing a management approach to reform the EU judiciary, arguing against an increase in the number of Judges at the GC, and appraising the reform approach taken by the European Court of Human Rights. The Dehousse papers are available under http://www.egmontinstitute.be, under the label of 'Egmont Papers' (last accessed on August 14, 2018).

competence of the candidate, understood as his knowledge of EU law, of a number of European languages, and his management skills. Second, the criterion of stability, emphasized by the President as paramount for the good functioning of the General Court. It was very important that political reasons would not impede the renewal of a judge's mandate. It was notorious that judges need a certain amount of time to become fully operational once appointed and that they are generally less operational in the last months of their mandate. The third criterion was geographic balance."[240]

It becomes clear that way in which judicial competence was evaluated in the Skouris reforms continues to follow the individualist paradigm of judicial competence – knowledge of EU law, language skills, and management competence. Cognitive diversity does not feature on the list. According to the CJEU itself, the reforms were meant to improve the GC's performance on several factors all at once: "speed, quality, coherence and, in short, *authority* of its case-law."[241] As mentioned, the GC has grown to two Judges per Member State as from September 1, 2019.

There is skepticism as to the real efficiency gains of the Skouris reforms. Some bemoan the abandonment of a broader use of the single judge at the GC.[242] From the point of view of PEA, however, the doubling of judges and the increase in chamber size is *per se* not bad. Now, Dehousse criticizes the rejection of the use of specialized courts as it strengthens the Member States' role in the appointment of Judges following the principle

---

[240] Franklin Dehousse, *The Reform of the EU Courts (II): Abandoning the Management Approach by Doubling the General Court*, Egmont Paper 83, March 21, 2016, available under: http://www.egmontinstitute.be/content/uploads/2016/03/ep83.pdf.pdf?type=pdf (last accessed on: August 14, 2018), 20.

[241] CJEU, *2017 Annual Report: Judicial Activity*, 138, my ital.

[242] Laurent Coutron, "The Changes to the General Court", in: Emmanuel Guinchard and Marie-Pierre Granger (eds.), *The New EU Judiciary: An Analysis of Current Judicial Reforms*" (Alphen aan den Rijn: Kluwers Law International, 2018), 144–160, 144-146. He writes:

> "This reform runs the considerable risk of resulting in a fall in the Court's productivity. Unless the number of appeals brought before the General Court doubles over three years, the judges' workload will in fact melt away. These additional judicial posts may, however, allow the case law of the General Court to 'go upmarket', that is to say be handed down by court formations composed of greater numbers of judges and, consequently, result from more detailed, in-depth argument […]."*

* ibid., 144–145.

of equal representation and hinders the appointment of specialists for areas such as patent law.[243] But what may seem troubling from a managerial point of view can prove beneficial from a cognitive point of view. Clearly, where expertise outweighs diversity, as it might in some cases concerning patent law, the reforms are a missed opportunity. With a view on the varied docket of the GC, however, equal representation can indeed boost its epistemic reliability. While heavily criticized from parts of the European judiciary, under the aspect of epistemic reliability the Skouris reforms are not nearly as bad as the critics have claimed them to be.

Now, chamber size matters from the point of view of statistical aggregation, the domain of the other mechanism of collective wisdom discussed here besides the DTA, the CJT. If we assume all Judges to be competent in the sense of the CJT, then we can get an idea of the epistemic costs of smaller chamber sizes. While this paper avoids technicalities as much as possible, some explanation of Condorcetian competence is needed here. Amongst other conditions,[244] the classic CJT applies in a decision between two choices, one of which is correct. Hence the effort in Section II to find a binary task description for international adjudication. The decision-making group must make its choice by majority voting. Furthermore, the members of the decision-making group individually need to be competent. Under the CJT, competence means a probability of choosing the correct alternative anywhere above 0.5. Probabilities are usually expressed in a range between 0 and 1. In this context, 0 means certainty of choosing the wrong answer, 1 means certainty of choosing the correct answer, and 0.5 expresses a random chance of choosing the correct answer, just like the roll of a fair die. The CJT then states that with competent group members, the probability of the majority of the group to vote for the correct alternative approaches certainty (i.e., 1) as the number of voters in the group goes to infinity.

This also means that the greater the competence of the members of a group, the steeper the increase in the likelihood of the majority vote to be correct. While the

[243] Cf. Dehousse, "The Reform of the EU Courts (II)", op. cit. 240, 5.

[244] For an overview of the applicability conditions of the CJT, see Franz Dietrich and Kai Spiekermann, "Jury Theorems", Working Paper, August 1, 2016, available under: http://www.franzdietrich.net/Papers/DietrichSpiekermann-JuryTheorems.pdf (last accessed on: February 19, 2020); Kai Spiekermann and Robert E. Goodin, "Courts of Many Minds", in: (2012) 42 *British Journal of Political Science* 3, 555–571, 556;: Goodin and Spiekermann, *An Epistemic Theory of Democracy*, op. cit. 143, 17-18.

difference between a chamber of three and five Judges may be small (unless the members are extremely competent), the differences between a Grand Chamber of 15 and the smaller chambers can already be significant. According to the classical formulation of the CJT, the majority in a Full Court with 28 Judges would, if the individual judges are deemed to have a competence of 0.7, almost certainly reach the correct decision.

It is important to stress again that this is an *a priori* calculation which would only hold true if the applicability conditions of the classical CJT are met. Those familiar with the Theorem will be quick to object that for the CJT to work, the votes in the group must be statistically independent of each other. In our case, that means that the Judges' probabilities to vote for the correct decision have to be independent from each other. This is likely not the case empirically. Robert Goodin and Kai Spiekermann provide a list of common causes which can render people's votes dependent: a shared opinion leader; a shared ideology; shared psychological mechanisms such as heuristics and biases; fundamental shared properties such as common training[245] or a common social background; shared evidence, background information, or theories; other voters, which are followed directly by their peers.[246] However Goodin and Spiekermann's *An Epistemic Theory of Democracy*[247] is an attempt to weaken the assumption of the classic CJT and to come up with jury theorems that are applicable in the real world. There are two things to take away from this: First, it is still valuable for scholars of IC legitimacy to think about the classic CJT's tenets, given that weaker assumptions might be available to carry the results. Second, the importance of common causes demonstrates the connection between PEA's *a priori* benchmark – its comparative 'positive' institutional design agenda – and the importance of empirical behavioral research – Elster's 'negative' institutional design. This is not to say that the latter is the easier task of both. The CJEU's decision making is to large extends a black box as it does not publish dissenting opinions and its deliberations are secret.

---

[245] Adrian Vermeule, "Many-Minds Arguments in Legal Theory", in: (2009) 1 *Journal of Legal Analysis* 1, 1-44, 6.

[246] Goodin and Spiekermann, *An Epistemic Theory of Democracy*, op. cit., 143, 55. Some have thus argued that discussion and deliberation are detrimental to independence. Cf., for example, Rawls, *A Theory of Justice*, op. cit. 147, 538 and Bernard Grofman and Scott L. Feld, "Rousseau's General Will: A Condorcetian Perspective", in: (1988) 82 *American Political Science Review* 2, 567–576.

[247] Op. cit. 143.

We are also able to see now how the competence assumption and the independence assumption interplay. If the independence assumption is met but the competence assumption violated (i.e., the competence of the members of the group is below 0.5), then the CJT is put into reverse. The probability that a majority vote will be correct will converge to zero as the number of voters increases. Goodin and Spiekermann rightly assume that it is more likely that a group's confusion stems from one or more of the common causes above, rather than that they are all delusional in their own, idiosyncratic ways.[248] If we assume, for example, that the independence assumption is violated completely because all voters follow one incompetent opinion leader with a probability of voting correctly of 0.4, then the probability of the majority being correct does not converge to 0 but finds its floor in the individual probability of voting correctly of the opinion leader (i.e., 0.4).[249] In this example, the opinion leader is not competent and the majority's competence is statically the same, no matter how many people are in the group. The same is true if competent individuals follow a competent opinion leader with a probability of voting correctly of 0.6. Since the independence condition is violated, the majority's probability of being correct does not converge to 1 as the group size increases but finds its ceiling in the opinion leader's competence of 0.6.[250] Altogether, Goodin and Spiekermann help us to see that the independence assumption at its extremes of either being completely fulfilled or violated renders the group competence to either follow a straight line (if violated) or an S-curve (if fulfilled). The competence assumption then defines the level of the straight line or whether we follow the S-curve to its top or bottom part.[251]

Goodin and Spiekermann thus deviate from the classical formulation of the CJT. Among their recent work, their 'best responder corollary' is interesting for our purposes, not least because of the problem of expert evidence in international adjudication. The corollary will be introduced in the following sub-section, which deals with expert advice.

---

[248] Goodin and Spiekermann, *An Epistemic Theory of Democracy*, op. cit. 143, 55.
[249] Ibid., 58.
[250] Ibid.
[251] In this mapping, the x-axis depicts the individual competence without common causes (opinion leader), and the y-axis depicts the probability of a correct majority winner, see ibid., 58–59.

For now, we must concede that in real-world conditions, even a Grand Chamber will likely not converge to 1 (i.e., the truth) but at best to the probability that, given the state of the available evidence and other aspects of the decision situation, its best responding Judge would decide correctly, given that this Judge also acts as an opinion leader. Still, we may believe that by increasing the chamber size we are raising the likelihood that a Judge with a higher probability of deciding correctly is added to the group. Moreover, we can also conclude that small chambers likely benefit less from judgment aggregation than bigger chambers, as theorized by the CJT. Small chambers are also unlikely to benefit from cognitive diversity, as a smaller number of Judges is drawn from the same population as larger chambers are, thus lowering the probability of assembling a cognitively diverse group of Judges. Efficiency, as Advocate General Sharpston suspected in her critique of the Skouris reforms, may indeed come with a loss of quality. The increase of Judges at the GC, however, has little to do with it in our framework.

c. Recourse to Experts

Finally, let us think about ICs' use of court-appointed experts. We can consider such experts as an extension of the epistemic agent of interest which the court forms.[252] Not all ICs necessarily have the inquisitorial powers to appoint their own experts (and hence may rely on the adversarial evidence adduced by the parties). Should they have such powers, not all ICs make frequent use of them.[253] Some – like the International Court of Justice – have been reluctant to appoint experts. Others – like the World Trade Organization's Dispute Settlement system – have made frequent use of their powers.[254]

In theory, the appointment of *ad hoc* experts consists of two steps which are epistemically interesting: first, the selection of the expert by non-expert judges, second, the assessment of the expert's opinion by non-expert judges. Courts are likely to develop judicial practices regarding the identification of relevant experts, their consultation, and

---

[252] Cf. again, Dillon, op. cit. 16.

[253] As to the different traditions by which courts are handling evidence, see: Eric Barbier de La Serre and Anne-Lise Sibony, "Expert Evidence Before the Court", in: (2008) 45 *Common Market Law Review* 4, 941–985.

[254] Dunoff and Pollack, "International Judicial Performance and the Performance of International Courts", op. cit. 54, 273–275.

the interpretation of their reports.[255] Courts might prove to be reluctant to not follow the opinion of an expert they themselves appointed: The CJEU, for example, so far always follows the expert's advice once it appoints one, even against the objections from the European Commission.[256] Nevertheless, some voices in the literature[257] as well as appellants and the European Commission[258] argue that expert reports should be commissioned more frequently by the CJEU.[259] Others believe that it "undermined the effective requirement of judicial review" as it may increase the duration of the proceedings.[260] What due process requires here is of course not only a question of epistemic values, but also of transparency, fairness, and impartiality.[261] And due process values become especially crucial when the experts consulted are supposed to review scientific assessments they themselves were involved in making.[262]

In the institutional dimension of legitimacy, however, we can single out the epistemic question. Can we, under PEA, make an argument for the increased use of court-appointed experts? Here, the best responder corollary by Goodin and Spiekermann is of interest again. It is not implausible to think that those who are taking a vote in a decision-making group do not have direct access to the state of the world but only have access to evidence

---

[255] Cf. ibid., 274. The WTO panels are a good example for since they frequently consult experts, at a rate of nearly 20 percent, cf. ibid.

[256] Cf. de La Serre and Sibony, "Expert Evidence Before the Court", op. cit. 253; See Joined Cases C-89/85, C-104/85, C-114/85, C-116/85, C-117/85, and C-125/85 to C- 129/85, *A. Ahlström Osakeyhtiö and Others v. Commission* [1993] ECR I-01307.

[257] Cf. Damien Geradin and Nicolas Petit, "Judicial Review in European Union Competition Law: A Quantitative and Qualitative Assessment", Tilburg Law School Research Paper No. 01/2011 (October 26, 2010), available under: http://dx.doi.org/10.2139/ssrn.1698342 (last accessed on: September 10, 2018).

[258] Cf. Case C-413/06 P, *Bertelsmann and Sony Corporation of America v. Independent Music Publishers and Labels Association (Impala)* [2008] ECR I-04951, in which the appellants and the Commission claimed the GC overstepped the scope or review and substituted its own assessment for that of the Commission, "without proving the existence of manifest errors of assessment vitiating the contested decision and without asking for a report from an economic expert to be obtained", ibid., para 138.

[259] Others disagree, cf. Ian Forrester, "A Bush in Need of Pruning: The Luxuriant Growth of 'Light Judicial Review'", in: Claus-Dieter Ehlermann and Mel Marquis (eds.), *European Competition Law Annual 2009: The Evaluation of Evidence and its Judicial Review in Competition Cases* (Oxford: Hart, 2011), 407-452; Yves Botteman, "Mergers, Standard of Proof and Expert Economic Evidence", in: (2006) 2 *Journal of Competition Law and Economics* 1, 71-100.

[260] Marc Jaeger, "The Standard of Review in Competition Cases Involving Complex Economic Assessments: Towards the Marginalisation of the Marginal Review?", in: (2011) 2 *Journal of European Competition Law & Practice* 4, 295–314, 314.

[261] Cf. Dunoff and Pollack, "International Judicial Performance and the Performance of International Courts", op. cit. 54, 286.

[262] Ibid. The authors are referring to the case of *US/Canada – Continued Suspension*: Appellate Body Report, Canada – Continued Suspension of Obligations in the EC-Hormones Dispute, WT/DS321/AB/R (October 16, 2008), paras 477, 479, 481

– understood as a common cause in the sense stated above – about the state of the world.[263] Goodin and Spiekermann describe the plausibility of this setup by considering jurors in a criminal trial:

> "The state of interest is whether the defendant is innocent or guilty. However, this state of the world does not directly cause the votes to be in favour of conviction or acquittal. Rather, the jurors will try to infer the state of the world (guilt or innocence) by observing the evidence available. […] The important upshot of the example is that we hardly ever observe the state of the world directly. Instead we try to get at the state of the world by more indirect means, especially evidence and testimony."[264]

This then leads to a different interpretation of the CJT. Framed in our setting of judicial decision making this means that if observing evidence is the only way for judges to make inferences about the state of the world because the effect of state of the world on judges is completely mediated through evidence, then the CJT is not about tracking the truth (i.e., the state of the world) directly. Instead, it is about the ability of individual judges and judicial chambers to reach the same judgment as the best responder would on the basis of the available and shared evidence.[265] The CJT then shows that the majority of a large group of competent and independent (conditional on the evidence) voters converges toward the judgment of the best responder. The best responder is the voter who processes the evidence in the best possible way and who chooses the option most likely to be true given that evidence.[266] Note that 'voter' here denotes both judges and appointed experts. This is a special case of Goodin and Spiekermann's best responder corollary.[267]

---

[263] See further: Krishna K. Ladha, "The Condorcet Jury Theorem, Free Speech, and Correlated Votes", in: (1992) 36 *American Journal of Political Science* 3, 617–634; Krishna K. Ladha, "Condorcet's Jury Theorem in Light of de Finetti's Theorem", in: (1993) 10 *Social Choice and Welfare* 1, 69–85; Krishna K. Ladha, "Information Pooling through Majority-Rule Voting: Condorcet's Jury Theorem with Correlated Votes", in: (1995) 26 *Journal of Economic Behavior & Organization* 3, 353–372.

[264] Goodin and Spiekermann, *An Epistemic Theory of Democracy*, op. cit. 143, 71–72.

[265] Ibid.; See further: Franz Dietrich and Christian List, "A Model of Jury Decisions Where All Jurors Have the Same Evidence", in: (2004) 142 *Synthese* 2, 175–202.

[266] Goodin and Spiekermann, *An Epistemic Theory of Democracy*, op. cit. 143, 72–73.

[267] Ibid., 77. The best responder corollary has implications for the laypeople-experts problem, which lies at the heart of the epistemic problems of expert evidence in court. The best responder is someone who does the best epistemically that can be done, given the decision situation. The decision situation is determined by common causes which can either be truth-conducive or misleading and cannot be changed by the

Under PEA, a more frequent use of an IC's inquisitorial powers would only be advocated for if it indeed increased the IC's epistemic reliability. If the IC follows the expert's opinion, we can technically understand experts here as opinion leaders. Experts as opinion leaders can then either provide a floor or a ceiling of epistemic reliability. The latter would only be acceptable if the Judges as a group cannot surpass the expert's individual reliability. In collective-choice contexts, epistemic deference to experts is thus not necessarily a dominant strategy. Consider the following example:

> "Imagine a set of 101 individuals, all of whom are independent of one another. Suppose that each of them correctly assesses the likelihood that he will be right on the question under discussion as […] 0.55; and each of the correctly assesses the likelihood that the one expert available to them will be right on that question as […] 0.70. If each of them defers to the expert, adopting the expert's conclusion as his or her own, then the probability the majority among them will be right is just the probability that the expert is right, 0.70. But if each of them disregards the expert and votes on the basis of his or her own assessment, the probability that the majority among the will be right is 0.84."[268]

While these conditions do not neatly match our decisional context, the example nevertheless illuminates the power of numbers behind a collectivist approach to competence. If judges are better than random in choosing between two fact-sensitive alternatives – for example whether to ban or not ban a certain chemical substance as in

---

responder, cf. ibid. If the situation is misleading, even the best responder will vote incorrectly. Voter competence is then understood as the probability to track the vote of the best responder, cf. ibid, 78.

As the technicalities of the best responder corollary go beyond what can be discussed here, the following proposition by Goodin and Spiekermann must suffice:

> "If the votes are independent conditional on the situation and the voters are best-responder trackers, then the probability of a majority vote being correct converges to […] the probability that the decision situation is truth-conducive, as the group size increases."*

* ibid. The corollary highlights two potential sources of differences between experts and laypersons. First, experts might simply be closer to "replicating the judgment of the best responder to any given decision situation". Second, experts may face a better, since more truth-conducive, decision situation than laypersons. In the first way, experts are better at interpreting the same body of evidence shared with the laypersons. In the second way, experts have a larger and more reliable body of evidence available for interpretation, cf. ibid., 81.

[268] Ibid, 319.

Methanex – then, as a group, they may achieve satisfying rates of accuracy when voting on the evidence brought to the court by the parties. Of course, where the available adversarial evidence is insufficient to base a decision on, as it was the case in *Afton Chemical*, then the judges may resort to appointing an expert with access to expert evidence. In the second case, the decisive epistemic question then moves from assessing the evidence to choosing the right expert which can satisfy our demands to act as a reliability-increasing opinion leader.

Admittedly, this is not a conclusive answer as to whether or not ICs should more frequently appoint their own experts. However, the calculatory example by Goodin and Spiekermann shows that if reliable expertise is not readily available, judges as a group of non-expert, but better-than-random judges do not necessarily fare epistemically bad. The example also shows that smaller chambers are more likely to benefit from commissioning an expert's report and treating the expert as an opinion leader than bigger chambers (of minimally competent judges). Interestingly, smaller chambers are actually more likely to appoint an expert. The cases in which the CJEU commissioned an expert's report are too few to compose a solid sample for a statistical analysis. This does not mean, however, that there are no recognizable patterns. If we sort out those cases from the sample in which the same chamber decided over a similar issue without joining those cases together,[269] we are left with 21 decisions in which the Court itself appointed an expert. Of those 21 decisions to commission an expert's report, 12 were made by a three-judge chamber, 6 were made by a five-judge chamber, two were made by a seven-judge chamber and only one was made by a 13-judge chamber. Now, due to the extremely small sample size we can only speculate about the causes. For example, in Grand Chamber cases the parties may submit better adversarial expertise, thus making the need for additional inquisitorial expertise obsolete.

---

[269] Those are the so-called Dyestuff cases which were not joined together, cf.: Case 48/69, *Imperial Chemical Industries Ltd. v. Commission* [1972] ECR 00619; Case 49/69, *Badische Anilin- & Soda-Fabrik v. Commission* [1972] ECR 00713; Case 51/69, *Farbenfabriken Bayer AG v. Commission* [1972] ECR 00745; Case 52/69, *J. R. Geigy AG v. Commission* [1972] ECR 00787; Case 53/69, *Sandoz AG v. Commission* [1972] ECR 00845; Case 54/69, *SA française des matières colorantes (Francolor) v. Commission* [1972] ECR 00851; Case 55/69, *Cassella Farbwerke Mainkur AG v. Commission*, ECR 887; Case 57/69, *Azienda Colori Nazionali - ACNA S.p.A. v. Commission* [1972] ECR 00933.

d. Interim Conclusion and Behavioral Implications

Applying PEA to the design of ICs reveals the value of adding judges, especially ones with diverse ways of thinking, to the bench. Their diverse cognitive models may even trade off against their expertise. Psychological research on economic prediction making suggests as much. Adding diverse economic thinkers[270] to a first prediction maker improves the accuracy of the group's prediction as the group size increases. It does so even if the added models are worse than the first model and the returns on accuracy diminish with each new thinker added.[271] Thus, adding non-lawyers to ICs benches likely helps increase their reliability. Non-lawyers as opinion-leading experts, however, only help increase reliability if they are highly competent, especially if the judicial chamber in the case at hand is a larger one. Under uncertainty, there may be cases in which we are epistemically better off if we put an economist on the bench as an equal amongst equals versus appointing her as an expert who acts as an opinion leader.

As mentioned, the comparative, *a priori* benchmark of truth-tracking requires behavioral empirical input. For example, whether or not the international composition of IC courts leads to epistemically better results also depends on whether or not IC judges become biased and tend to allocate decision making authority towards the international (or supranational) level, away from domestic jurisdictions.

After all, as Vermeule writes, the epistemic quality of the laws is only one value of many against which it trades off.[272] Courts are better off with *some* extra-legal knowledge than with none.[273] Institutional design under PEA is thus about optimizing, not maximizing truth-tracking.

---

[270] Diverse means here that they draw on different predictive models.

[271] See: Albert E. Mannes, Jack B. Soll and Richard P. Larrick, "The Wisdom of Select Crowds", in: (2014) 107 *Journal of Personality and Social Psychology* 2, 276–299.

[272] See also: Vermeule, "Many-Minds Arguments in Legal Theory", op. cit. 245.

[273] "On the case level, courts might prefer to navigate with an inaccurate map than with no map at all", Jack Goldsmith and Adrian Vermeule, "Empirical Methodology and Legal Scholarship", in: (2002) 69 *University of Chicago Law Review* 1, 153–167. Cf. the argument made for the empirical legal studies movement in: Theodore Eisenberg, "The Origins, Nature, and Promise of Empirical Legal Studies and a Response to Concerns", in: (2011) *University of Illinois Law Review* 5, 1713–1738, 1720.

2. Judicial Reasoning

As mentioned before, the institutional redesign of ICs would almost always require treaty change. Recommendations for institutional design are therefore uncertain to be implemented. Another option for institutional designers to suggest improvements in institutional reliability is to propose reasoning principles for judges, leaving it essentially to the judiciary to adopt those suggestions. This move has been theorized by 'contextual institutionalism'[274]. It argues that judges should contextualize the issue at hand and "consider institutional factors when attributing weight to the views of other officials or the threat of uncertainty"[275]. This implies the view that judicial discretion can be structured by the use of principles of restraint and deference and that fallible judges can indeed be trusted enough to balance these considerations satisfactorily.[276] How likely IC judges are to contextualize the issue at hand institutionally likely depends on context. Threats to an ICs judicial authority are likely to lower judges' willingness to consider institutional factors.[277]

Principles of judicial reasoning, such as principles of deference or restraint, are a key concern for institutional approaches to judicial authority. To oppose their plausibility would mean to hold judges to be infallible.[278] On the other side of the spectrum, restrictive accounts of judicial authority propose an absolute supremacy of legislatures and technocrats through bright-line rules of deference. Their belief in the omnipresence of judicial fallibility and uncertainty is so strong that judicial competence appears too low to entrust judges with anything but clear textual interpretation.[279] A satisfying answer can

---

[274] Jeff A. King, "Institutional Approaches to Judicial Restraint", in: (2008) 28 *Oxford Journal of Legal Studies* 3, 409–441, 430.

[275] Ibid.

[276] Ibid.

[277] For example, the CJEU has proven to be reluctant to draw directly on the jurisprudence of the ECtHR after the Charter entered into force (see further below).

[278] For an argument that comes close to an implicit assumption of judicial infallibility, see: T. R. S. Allan, "Human Rights and Judicial Review: A Critique of 'Due Deference'", in: (2006) 65 *Cambridge Law Journal* 3, 671–695.

[279] Cf. Eskridge's critique of Vermeule's *Judging under Uncertainty*, William N. Eskridge, Jr., "Review: No Frills Textualism", in: (2006) 119 *Harvard Law Review* 7, 2041–2075, 2066.

only be given in a comparative approach, allocating authority amongst all (fallible) institutions available to resolve a given problem.[280]

According to an analysis by Aileen Kavanagh, judges recognize their "institutional shortcomings" in three main situations: when there is a deficit of institutional competence, expertise, or institutional or democratic legitimacy.[281] One of the contributions PEA seeks to make is to develop principles which govern all three situations under a common framework which understands the judicial role in the constitutional system as a reliable detector of breaches of law.[282] PEA may also give an answer to the question of when judges can be legitimately *activist*, a feature few accounts of IC authority have. Take the democratic legitimacy of ICs as an example. Such concerns were traditionally discussed under the umbrella of state sovereignty which was said to be blind towards domestic democratic processes.[283] More recently, Andreas von Staden has suggested the idea of "normative subsidiarity", i.e. a "democratically informed standard of review that guides courts in the exercise of their review and dispute settlement functions."[284] Von Staden's review standard is thus sensitive towards the democratic legitimacy of the decision-making institution. Hence, it is comparative in nature and seeks to allocate decision-making authority towards the democratically more appropriate institution. But how could normative subsidiarity function as anything else but a principle justifying the restraint of courts? It does not say much about whether or when courts are justified in being *activist*. Moreover, what benchmark is to be applied to the actual judgments of courts which pass von Staden's subsidiarity test? A mere *ex-negativo*

---

[280] See, again in this regard: Neil K. Komesar, *Imperfect Alternatives: Choosing Institutions in Law, Economics, and Public Policy* (Chicago: University of Chicago Press, 1994); Neil K. Komesar, *Law's Limits: The Rule of Law and the Supply and Demand of Rights* (Cambridge: Cambridge University Press, 2001).

[281] Aileen Kavanagh, "Deference or Defiance?: The Limits of the Judicial Role in Constitutional Adjudication", in: Grant Huscroft (ed.), *Expounding the Constitution: Essays in Constitutional Theory* (New York: Cambridge University Press, 2008), 184–215, 192.

[282] Kavanagh also argues that principles of judicial restraint must start with an account of the judicial role and its limits in the constitutional framework, cf. ibid., 190.

[283] Von Bogdandy and Venzke, *In Whose Name?*, op. cit. 1, 198.

[284] Andreas von Staden, "The Democratic Legitimacy of Judicial Review Beyond the State: Normative Subsidiarity and Judicial Standards of Review", in: (2012) 10 *International Journal of Constitutional Law* 4, 1023–1049, 1026.

justification, claiming that all decisions which are not subject to normative subsidiarity are thus legitimate, appears to be an extremely thin standard of legitimacy.[285]

As we can see with von Staden's suggestion of normative subsidiarity, a mere *ex negativo*-argument is normatively not satisfying. The absence of the need for deference does not legitimize *any* kind of activism. Principles of judicial restraint or deference thus require complementary principles of activism. The principles suggested below require further calibration than can be provided here. The aim of this paper is merely to present a concretization of the framework of PEA in order to provide further ground for a debate about its merits. Before we begin to draft our principles, we need to clarify our terminology and provide a simple model of the decisional situation.

a. Modeling the Decisional Situation

Let us first clarify some of the assumptions we are making about the decisional situation. The starting point is the observation made by Vermeule, that other than scientists, who can leave unanswered questions to another research study, judges are forced to decide the issue at hand *now.*[286] Furthermore, in the institutional context we are considering here, there usually already exists a judgment (or opinion) of another institution. The cases we are considering are thus regularly second opinions, whether on international or national institutions.[287] Obviously, our model then only catches a subset of the cases which are decided before ICs, namely cases of review. Here, judicial authority fulfills a compensatory role: it corrects some of the relative unreliability and other agential shortcomings of other institutions.[288] In some cases, however, judges will become activist

[285] From the point of view of democratic legitimacy, however, the greatest interest lies exactly with cases of judicial activism, cf. Svenja Ahlhaus and Peter Niesen, "Der Axtmörder und die Menschenrechtsgerichtsbarkeit: Überlegungen zum Fall Hirst des Europäischen Gerichtshofs für Menschenrechte", in: Claudio Franzius, Franz C. Mayer, and Jürgen Neyer (eds.): *Grenzen der europäischen Integration: Herausforderungen für Recht und Politik* (Baden-Baden: Nomos, 2014), 149–172, 159. It further seems that the deliberative democratic properties of international and supranational courts are not enough to justify their de facto authority, cf. ibid., 156–164.

[286] Cf. Vermeule, *Judging under Uncertainty*, op. cit. 12, 3.

[287] For a similar focus, see Adrian Vermeule, "Second Opinions and Institutional Design", in: (2011) 97 *Virginia Law Review* 6, 1435–1474.

[288] For an account of the compensatory role of general authority, see: Daniel Viehoff, "Authority and Expertise", in: (2016) 24 *The Journal of Political Philosophy* 4, 406–426.

and provide a first opinion on a given matter in the case at hand. This will most often be the case when judges identify a normative gap in the legal order which requires closing.

As regards second opinions, judges may exercise judicial restraint or deference, thus limiting the intensity of their review. For the purposes of this paper, it suggests itself to introduce an analytical distinction between the two modes. So far, we have used the notions of restraint and deference somewhat interchangeably. From now on, let us define judicial restraint as the case in which judges refrain from making a judgment on the reliability of the first opinion by another institution. Let us then define epistemic deference as the case in which judges judge the other institution's first opinion on a given matter to be more reliable than their own. Other than in cases of judicial restraint, the first opinion thus becomes part of the judicial second opinion, without deliberation or discussion of its merits. Finally, judicial activism is the case in which judges hold their own reliability to be at least sufficient enough to form a first opinion on a given matter.

Lastly, we are again going to follow our observations in Section II and focus on cases under empirical and normative uncertainty. These cases will be discussed under three types of epistemic deference, differentiated by the type of institution they concern: democratic (i.e., parliamentary), judicial, and technocratic deference.

Obviously, as it is the case with principles of restraint, the application of principles of deference lies within the discretion of judges. We may have doubts about the reliability of IC judges to correctly identify the situations in which deference is due. However, judges may err from time to time without risking their authority to become illegitimate with every single transgression. PEA is only interested in systematic judicial behavior, demonstrated over a sufficiently large set of decisions.[289]

b. An Epistemic Principle of Democratic Deference

Let us begin by drafting an Epistemic Principle of Democratic Deference. PEA does not consider public attitudes towards normative principles as a means to enhance democratic

[289] This raises the obvious question as to when a set of decisions is sufficiently big enough. Here, both quantitative as well as qualitative aspects should be taken into account. Errors about jurisdiction appear qualitatively particularly grave. A general reluctance to appoint experts may reach a quantitative threshold.

legitimacy.[290] Rather, such a principle would follow the compensatory function of judicial authority and establish an epistemic best responder standard (see above). Several formulations of such a principle can be offered, each with increasing feasibility towards institutional application.

In its ideal version, the Epistemic Principle of Democratic Deference first and foremost expresses the compensatory function of second opinions. It requires that the expertise, the evidentiary system, the access to better information or any other epistemic benefit which being in a position of exercising public authority to issue second opinions is supposed to provide should amount to something when compared to the epistemic reliability of the public at large, i.e., the citizens over which the authority is meant to rule.

> *An Ideal Epistemic Principle of Democratic Deference*
> The epistemic reliability of a public authority over domain D must not be lower than the epistemic reliability of the public at large over domain D.

To measure the level of epistemic reliability of the public at large would require aggregating citizens individual judgments, i.e., their votes and not merely their preferences as expressed, for example, in opinion polls.[291] As much as the ideal principle rests on the classic CJT, the majority of the public at large, powered by the law of large numbers, would be nearly infallible. The upper bound of the benchmark of epistemic reliability would thus be truth for all questions (or tasks) that are truth-apt and fall into domain D.

This, however, is highly unlikely. As shown above, for the classic CJT to work the votes must be statistically independent. People's opinions are likely influenced by common causes. Such common causes may introduce dependence on (1) a shared opinion leader; (2) a shared ideology; (3) shared psychological mechanisms such as heuristics and biases; (4) fundamental shared properties such as common training or a common social background; (5) shared evidence, background information, or theories; (6) other voters, which are followed directly by their peers.[292]

---

[290] For such a principle of democratic restraint, cf. Alice Baderin, "Political Theory and Public Opinion: Against Democratic Restraint", in: (2016) 15 *Politics, Philosophy & Economics* 3, 209–233.

[291] See the illuminating discussion (and rejection) of the role of public opinion for political theory in ibid.

[292] Goodin and Spiekermann, *An Epistemic Theory of Democracy*, op. cit. 143, 55.

We may also run into a regress problem: the public at large is required to vote but who formulates and presents the alternatives to vote on and how to vote for and against them?[293] Also, an ideal principle runs the risk of violating the commitments of public reason: those who find themselves holding a minority view would be silenced by an infallible majority with a monopoly on truth claims. Dissent, however, can be epistemically beneficial.[294]

With Goodin and Spiekermann we can agree that the true state of the world is most often not accessible to voters. Taking their best responder corollary into account, we can formulate a conditional principle:

> *A Conditional Epistemic Principle of Democratic Deference*
> The epistemic reliability of a public authority over domain D must not be lower than the epistemic reliability of the best responder in domain D.

The conditional principle, however, only solves one problem for the price of creating another. It is able to account for the distorting effects of common causes and thus takes a huge step towards real-world feasibility. The benchmark of epistemic reliability has moved from truth to truth-tracking. But now we are running the risk of establishing an expertocracy in domain D. We have no *a priori* reason to believe that the public at large is going to be the best responder and not some unaccountable eminence who happens to be the smartest person in a crowd. This conflicts with demands of procedural fairness we have for public authorities. It thus appears necessary to stage a democratic intervention and tilt the balance towards institutions with a higher rate of satisfying proceduralist demands.

Finally, we arrive at the most feasible – from the point of view of practical implementation *and* normative acceptability – version of the principle:

> *An Institutional Epistemic Principle of Democratic Deference*

[293] Josiah Ober, however, has suggested a model of "relevant expertise aggregation" in which domain experts play an important role in weighing and ranking options subject to judgment by ordinary citizens, cf. Josiah Ober, "Democracy's Wisdom: An Aristotelian Middle Way for Collective Judgment", in: (2013) 107 *American Political Science Review* 1, 104–122.

[294] Cf. Cass R. Sunstein, *Why Societies Need Dissent* (Cambridge, MA: Harvard University Press, 2003).

> The epistemic reliability of a public authority in domain D must not be lower than the epistemic reliability of the democratically most justifiable institution in system S over domain D.

In its institutional version, the Epistemic Principle of Democratic Deference is fully comparative while its benchmark of epistemic reliability remains committed to the standard of truth-tracking. As much as judicial review is supposed to, for example, provide an epistemically beneficial second opinion on legislative decisions and to correct some of its mistakes,[295] the principle in its institutional version needs to be fulfilled. The standard of the 'democratically most justifiable institution' is a deliberately open formula in order to not rest the legitimation of public authority only (or even dominantly) on epistemic reasons.

Let us now place the Institutional Epistemic Principle of Democratic Deference in a concrete institutional context. For a lack of parliamentary structures in the international level outside of the EU, let us focus again on the CJEU. The Principle would actualize itself whenever the European Parliament has been involved in the legal provision under review. Member State parliaments might even provide higher levels of democratic justification than the European Parliament. However, they cannot be considered in the confined space of this paper. Our concretized principle can thus be formulated as follows:

> *The Epistemic Principle of Democratic Deference in the EU*
> In cases of legislative review, the CJEU's epistemic reliability must not be lower than that of the European Parliament if it seeks to change the legal status quo.

Because of its numerosity and cognitive diversity, in an *a priori* assessment of its ability to decide between two well-defined options (CJT) or to find better solutions to a given problem (DTA), the European Parliament will score relatively high in terms of epistemic reliability. The European Parliament is currently composed of 751 Members of the European Parliament (MEPs).[296] Above we saw how swiftly collective competence

[295] Again, this is the tradition in which John Hart Ely's *Democracy and Distrust: A Theory of Judicial Review* (Cambridge, MA: Harvard University Press, 1981) stands out large. Vermeule seems to follow this tradition as well, see Vermeule, "Second Opinions and Institutional Design", op cit. 287.

[296] Cf. http://www.europarl.europa.eu/meps/en/home (last accessed on: December 10, 2018).

goes up under the classic CJT as the group size increases.[297] A simple majority of 376 MEPs would be enough to approach certainty if the MEPs individual likelihood of deciding correctly is above 0.50. The EU legislative process is of course more complex than that. In the so called ordinary legislative procedure, the European Parliament and the Council of the EU have equal say. In this procedure the Commission submits a legislative proposal to the Parliament and the Council who then must agree for it to become EU law. The Council by itself aggregates the votes of currently 28 Member State representatives and is thus only limitedly benefitting from numerosity when compared to the CJEU. The European Parliament, however, does benefit from its numerosity.

The European Parliament would fare similarly well under the lens of perspectival aggregation through cognitive diversity (following the DTA). Moreover, as has been argued by Krishna Ladha[298] already for statistical aggregation under the CJT, adding members to the group whose competence is worse than random can improve overall group performance, if their biases are negatively correlated to a sufficient degree with those of experts.[299] Under plausible (Condorcetian) conditions, a professionally diverse group may outperform a group of specialists. While professional diversity is argued to correlate with epistemic diversity, common professional training is a major source of correlated biases. The average individual competence may be pushed up by common training, but so are common blind spots and the risk of groupthink.[300] (American) lawyers have been argued to be particularly prone to share distinctive traits, compared to both the general population and other professionals.[301] It does not require much descriptive analysis to state that the professional training of the MEPs is more diverse than that of the judges of the CJEU. We must not forget that if the members of the group are all incompetent (i.e., worse than random), then the CJT is put into reverse and the likelihood of reaching the correct decision approaches 0 as the group size increases.[302]

[297] Hence, epistemic democrats do not need to draw on the entire electorate to arrive at the near-certain infallibility of the majority (given the people in the electorate are better than random). Representative democracy suffices (now given that the representatives in parliament are better than random).

[298] Ladha, "The Condorcet Jury Theorem, Free Speech, and Correlated Votes", op cit. 263.

[299] Vermeule, "Collective Wisdom and Institutional Design", op. cit. 12, 354.

[300] Ibid.

[301] Susan Daicoff, *Lawyer, Know Thyself: A Psychological Analysis of Personality Strengths and Weaknesses* (Washington, DC: American Psychological Association Books, 2004).

[302] See further: Goodin and Spiekermann, *An Epistemic Theory of Democracy*, op. cit. 143, 51.

Since the judges do not have direct access to the levels of competence of the MEPs, let alone of their own, the correct application of the Epistemic Principle of Democratic Deference would require judges to make fairly crude assessments of epistemic reliability. However, the aim of the framework of PEA is not to provide methods by which to decide cases mathematically but rather to give judges institutional grounds to trust parliamentary decisions. PEA therefore depends on a certain state of democracy's well-being, which might not be given in all historical constellations, as the current rise in populism demonstrates.

That said, if we assume our institutions to be populated by minimally competent (i.e., better than random) people, parliaments have epistemic advantages over courts which should matter in the institutional dimension of legitimacy. Let a stylized argument illustrate the point: Assume we are dealing with an important constitutional question as to whether a Treaty provision means A rather than B, which the CJEU thus decides in its Grand Chamber of 15 Judges. If those 15 judges are competent, they achieve a very high likelihood of reaching the correct decision, at least when measured in the classic understanding of the CJT.[303] However, a simple majority of 376 MEPs would be enough to surpass the collective competence of the Grand Chamber, even if the individual competence of MEPs was lower than that of the judges in the Grand Chamber, as long as it is not too low (i.e., still better than random).[304] Even when following the best responder

[303] If, for example, the average individual competence of the Judges on this question is 0.70 then their collective competence as a group of 15 is well above 0.90.

[304] However, the European Parliament may itself defer its judgment to its Committee on Legal Affairs, which currently has 25 members and is among the smallest of all European Parliament committees, Cf. Giulio Sabbati, *European Parliament: Facts and Figures*, April 2018, available under: http://www.europarl.europa.eu/RegData/etudes/BRIE/2018/614733/EPRS_BRI(2018)614733_EN.pdf (last accessed on: October 10, 2018), 12. The Committee takes over 'horizontal' as well as 'institutional' responsibilities besides its legislative activities. Under horizontal responsibilities fall tasks which relate to the interpretation and application of EU law; amongst its institutional responsibilities are tasks which relate to the choice of the correct legal basis for Union acts, cf. European Parliament, Committee on Legal Affairs, *Stocktaking of parliamentary committee activities during the 7th legislature*, August 12, 2014, available under: http://www.europarl.europa.eu/cmsdata/54165/att_20141103ATT92392-5641366759106240678.pdf (last accessed on: October 9, 2018), 3.

Now, if only lawyers were to self-select into the Committee, the legislature's benefit of cognitive diversity would disappear. However, this does not seem to be the case. (Not all members of the Committee on Legal Affairs provide public information on their CVs. Two of the Committee's vice chairs, however, do not have formal legal training. One is an economist and statistician, another one is a philologist and engineer, cf. http://www.europarl.europa.eu/meps/en/28377/LIDIA+JOANNA_GERINGER+DE+OEDENBERG_cv.html (last accessed on: October 10, 2018) and http://www.europarl.europa.eu/meps/en/124776/MADY_DELVAUX_cv.html (last accessed on: October 10, 2018).) Before we said that such committees can be epistemically beneficial if their suggested solution

corollary of Goodin and Spiekermann and not the classic CJT, the comparison should point in the same direction and locate the higher epistemic reliability with the Parliament. Now, one may argue that if the European Parliament decides correctly with near and the CJEU is lagging behind only by a small margin, the whole comparison is superfluous. Moreover, if both institutions are highly likely to be correct, how would there ever be a situation in which the CJEU believed the Treaty provision to mean A and the Parliament believed the same provision to mean B. Would they not have the same understanding if they both are nearly infallible?

It is important to not let the stylized examples misrepresent the argument made here. As regards the institutional legitimacy of second opinions, the framework of PEA can show that even if the individual competence of MEPs is lower (but not too low) than that of individual Judges, they *can* still collectively exceed the collective competence of the Judges as a group. And as much as the CJEU's legitimate authority is built upon its compensatory function in cases of second opinions, it requires the judges to be regularly epistemically more reliable than the institution under review.[305]

Obviously, none of this prevents the legislature from making mistakes. Hence the compensatory role for judicial review. In the domain of cases under technical complexity or scientific uncertainty, it would be very hard for the CJEU to change the legislative status quo unless the EU legislature would have committed a fairly obvious blunder. In other domains, such as fundamental rights reasoning where the interests of a minority could be at stake, the CJEU may prove to be more reliable than the Parliament. What is required are thus domain-specific considerations.

Afton Chemical is a good illustration of how the CJEU can draw on the Epistemic Principle of Democratic Restraint to decide cases under scientific uncertainty. As shown

---

gets voted on by the plenary. Another option is to model these committees as opinion leaders, cf. Vermeule, "Collective Wisdom and Institutional Design", op. cit. 12, 354. Vermeule seems to assess legislative committees as high-competence opinion leaders. In this line of reasoning, the CJT seems unable to carry the weight of the argument, at least in our example (25 MEPs versus 15 judges), and much of the work has to be done by the DTA, accounting for the higher cognitive diversity in the committee when compared to the CJEU. However, such committee deliberation can be epistemically beneficial as long as the legislature as a whole gets to vote on the committee's recommendations, cf. Goodin and Spiekermann, *An Epistemic Theory of Democracy*, op. cit. 143, 317.

[305] If judicial review is justified on grounds of a however-formulated right to be heard, then the reliability of the Court may not matter. But this understanding would not contribute to answering our question of whether the Treaty provision means A or B.

above, the judges had to review a directive which was based on the precautionary principle and established limits to the use of the chemical MMT. The producer of MMT, Afton Chemical, claimed those limits were not based on scientific evidence.

Typically for the European legal order, the problem was solved via a proportionality test. The CJEU thus asked whether the "European Union legislature attempted to achieve a degree of balance between […] the protection of health, environmental protection and consumer protection and […] the economic interests of traders […]."[306] At the heart of the case thus lied a dispute over the intensity of interference with health, environmental, and consumer concerns caused by the uncertain harmfulness of MMT. Without an epistemic benchmark of legitimacy in the background, scholars may easily subscribe to the view that in such cases judges should balance themselves different degrees of reliability of the epistemic premises involved and thus limit the legislature's discretion.[307] The Epistemic Principle of Democratic Deference points the other way and suggests that judges hold the asserted inferences of the legislature to be superior to their own and incorporate them into their own judgment. In *Afton Chemical*, the CJEU did exactly that, when it took over the intensity of interference assumed by the European legislator into their proportionality test. The judges might not have done this for epistemic reasons, but we can rationally reconstruct the decision by the CJEU as an act of epistemic deference. The CJEU did not fill the epistemic variables of the proportionality test with their own appreciation of the facts. From the point of view of PEA, this democratic deference is legitimacy enhancing.

However, our analysis of *Afton Chemical* also suggested that if there had been existing scientific research on the harmfulness of MMT that did not stem from industry but from public institutions, the judges might have been tempted to provide their own assessment of it. However, the Epistemic Principle of Democratic Deference would, all else being equal, favor a solution like the one found by the NAFTA Tribunal in *Methanex*: simply defer to the legislature's assessment of the scientific evidence and not form a judicial second opinion on its merits.

What about cases under normative uncertainty? After all, judges are experts in normative (at least legal) reasoning. Parliaments' institutional *raison d'être*, on the other

---

[306] *Afton Chemical*, op. cit. 91, para 56.

[307] For such a suggestion, see: Matthias Klatt and Johannes Schmidt, "Epistemic Discretion in Constitutional Law", in: (2012) 10 *International Journal of Constitutional Law* 1, 69–105.

side, seems not to lie in any domain specific expertise. Whether or not epistemic democratic deference is due in cases under normative uncertainty appears to be dependent on the legal method by which the normative question at hand is supposed to be answered. Vermeule, for example, argues at book length that in U.S.-constitutional cases in which the original intent of the Framers' of the American Constitution is to be determined, the American legislature has better institutional capacities to answer this historical question correctly than the U.S. Supreme Court. The latter's task is thus reduced to providing mere textualism.[308] It is beyond the means of this paper to provide an account of the legal methods deployed by ICs. Moreover, not all legal methods lend themselves to be applied by legislatures instead of courts. Above, we considered the proportionality test in *Google Spain* in which four interests had to be balanced. The case-specific balancing outcome could not have been anticipated by the European legislature (or by the Treaties themselves). However, the judicial activism of the CJEU in *Google Spain* to create what would later be called the 'right to be forgotten' (or the 'right to erasure') still lends itself to an institutional analysis, as our considerations below will show. The following two principles concern institutions with clear domain-specific expertise: courts and technocratic institutions.

c. An Epistemic Principle of Judicial Deference

The jurisdictions of ICs may overlap with each other and may infringe the authority of national courts. It is thus not unusual for ICs to answer legal questions which fall in the domain of expertise (and potentially the jurisdiction) of other courts. Is it possible to formulate a principle for when epistemic judicial deference is due based on the mechanisms of collective wisdom?

As argued above, under PEA we are determining institutional reliability first in an *a priori* approximation, drawing on the outcomes of comparisons under the mechanisms of collective wisdom. The goal of this *a priori* step is to give judges reasons to contextualize the issue at hand and to make an institutionally informed judgment as to when to defer to the judgments of other institutions and when not to. However, given the

[308] Cf. again, Eskridge, "Review: No Frills Textualism", op. cit. 279.

relative similarity of the institutional design of ICs and at least apex courts on the national level, the mechanisms of collective wisdom might not be informative enough for this task. For example, comparing the *a priori* reliability of a five-judge chamber of court A with the *a priori* reliability of the seven-court chamber of court B based entirely on the classic CJT and hence arguing for deference of A to B[309] seems way off the mark. Different approaches to judicial selection may produce differences in competence and, more interestingly, in cognitive diversity. However, for those ICs whose benches are filled exclusively with lawyers even judicial selection does not give us much as regards the application of the DTA.[310]

There is another mechanism of collective wisdom we have not yet discussed in this paper, the so called many-minds arguments of evolutionary aggregation. While the statistical and the perspectival model of aggregation both compare crowds and individuals to decide at the same point in time, the judicial practice of precedent and 'invisible-hand' mechanisms of legal evolution, to the contrary, both draw from (or compare) decisions over time.[311] As evolutionary aggregation through precedent only applies to judicial institutions, it seems appropriate to introduce it in the context of the Epistemic Principle of Judicial Deference.

Evolutionary models differ from statistical and cognitive models in that they do not pre-suppose any intentionality of the decision-makers to arrive at the correct decision.[312] Collective wisdom arises from human action but not by human design.[313] In his *A*

---

[309] The simplistic reasoning would go as follows: Given individual competence being equally above 0.50 and all else being equal, 7 is a higher number than 5, therefore the court with 7 judges is more reliable than the court with 5 judges.

[310] For example, von Bogdandy and Krenn argue that the Council of Europe has the "better" judicial selection process compared to the EU judiciary. 'Better' means here being based on a list of criteria which has been developed by representative institutions such as the Council of Europe's organ of the Committee of Ministers and the Parliamentary Assembly, cf. von Bogdandy and Krenn, "On the Democratic Legitimacy of Europe's Judges", op. cit. 188, 172. However, PEA's inherent instrumentalism is blind towards representativeness as much as it does not translate into greater cognitive diversity.

[311] Vermeule, "Collective Wisdom and Institutional Design", op. cit. 12, 346.

[312] Ibid.

[313] Vermeule cites Adam Ferguson, *An Essay on the History of Civil Society*, (New York: Georg Olms, 2000), 187. See further: Robert Nozick, "Invisible-Hand Explanations", in: (1994) 84 *American Economic Review* 2, 314–318; Edna Ullmann-Margalit, "The Invisible Hand and the Cunning of Reason", in: (1997) 64 *Social Research* 2, 181–198. These evolutionary models stem from the law and economics approach and have been brought forward for the common law which is supposed to evolve to efficiency through adjudication as well as for legislatures moving towards efficiency through interest groups, cf. Paul H. Rubin, "Why Is the Common Law Efficient?", in: (1977) 6 *Journal of Legal Studies* 1, 51–63 and Gary Becker, "Public Policy, Pressure Groups, and Dead Weight Loss", in: 28 *Journal of Public Economics* 3, 329–347.

Constitution of Many Minds, Cass Sunstein argues (from an American perspective) that the major approaches of constitutional interpretation all rely on some variation of a "many-minds" argument.[314] Sunstein's book is not without relevance for the IC context of this paper. It will be read here through the eyes of a discussion paper by Spiekermann and Goodin, who assess Sunstein's claims through the lens of the CJT.[315] Sunstein identifies three major approaches of constitutional interpretation:

> "*Traditionalists* insist that if members of a society have long accepted a certain practice, courts should be reluctant to disturb that practice. […]
> *Populists* believe that if most people accept a certain fact or value, judges should show a degree of humility – and respect their view in the face of reasonable doubt. […]
> *Cosmopolitans* believe that if many nations, or many democratic nations, reject a practice, or accept a practice, the US Supreme Court should pay respectful attention."[316]

Now, Sunstein claims that all three approaches rest on a many-minds argument of the kind that "if many people think something, their view is entitled to consideration and respect."[317] The respect does not stem from courtesy, but from the likelihood of larger groups being more likely to be right.[318] Sunstein further claims that "[t]he structure of the central argument is identical in all three contexts"[319]: the formal foundation is the CJT.[320] Sunstein then adjudicates amongst the three approaches on grounds of them sharing a common formal basis.[321] Conclusively, Sunstein ranks traditionalism as the epistemically

---

Both models only work if stakes are not asymmetrically distributed across groups. Otherwise, rent-seeking groups impose welfare losses both in legislation and in court, cf. Vermeule, "Collective Wisdom and Institutional Design", 347. See further: Einer R. Elhauge, "Does Interest Group Theory Justify More Intrusive Judicial Review?", in: (1991) 101 *Yale Law Journal* 1, 31–110.

[314] Sunstein, *A Constitution of Many Minds*, op. cit. 144.

[315] Spiekermann and Goodin, "Courts of Many Minds", op. cit. 244.

[316] Sunstein, *A Constitution of Many Minds*, op. cit. 144, ix–x, my ital.

[317] Ibid., ix.

[318] Spiekermann and Goodin, "Courts of Many Minds", op. cit. 244, 555.

[319] Sunstein, *A Constitution of Many Minds*, op. cit. 244, x.

[320] Ibid., 8–10.

[321] Spiekermann and Goodin remark that in his discussion of those issues, "Sunstein himself abjures formalism in favour of more context-sensitive lawyer-style discussions", Spiekermann and Goodin, "Courts of Many Minds", op. cit. 244, 555. As this thesis could be charged in a similar way, their results are an important lesson for the discussion of the applicability of the CJT.

strongest approach, followed by populism, and internationalism as the weakest.[322] In their formal analysis, Spiekermann and Goodin identify populism as plagued with fewer problems than traditionalism and cosmopolitanism.[323] They thus draw conclusions different from those of Sunstein.

As regards the Epistemic Principle of Judicial Deference, 'populism' is irrelevant. Let us therefore look at traditionalism and cosmopolitanism. First of all, the traditionalist and cosmopolitan setups are different from the original assumption of the CJT: the original CJT assumes that votes are cast either simultaneously or in ignorance or indifference to what other voters have done before.[324] In traditionalism and cosmopolitanism, courts make their decisions sequentially and "not only in knowledge of but also in deference to earlier courts' decisions."[325] The problem, however, is that if judges do not only follow their private signal as to what the correct answer in the issue at hand is but also the history of votes on the same matter, then instead of collective wisdom an informational cascade may emerge.[326] Informational cascades can generally occur where information is costly and it thus becomes rational to copy others' judgments.[327] In an informational cascade, the history of votes "constitutes such strong evidence in favor of one alternative that all judges will always follow the evidence of the history and never vote according to their own private signal."[328] This is epistemically bad because the information base on which future judgments are grounded can thus be very thin, if judges stop learning from their own signal and simply follow the historic voting record.[329]

---

[322] Sunstein, *A Constitution of Many Minds*, op. cit. 144, 211–214. For a different result as regards the cognitive value of recourse to recourse to international and comparative jurisprudence, see Jeremy Waldron, *Partly Laws Common to All Mankind: Foreign Law in American Courts* (New Haven: Yale University Press, 2012), Chapter Four.

[323] Spiekermann and Goodin, "Courts of Many Minds", op. cit. 244, 556.

[324] Ibid., 557.

[325] Ibid.

[326] Ibid., 558. Vermeule calls this the "Burkean Paradox", cf. Adrian Vermeule, *Law and the Limits of Reason* (Oxford: Oxford University Press, 2009), 75–77.

[327] See generally: Sushil Bikhchandani, David Hirshleifer, and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades", in: (1992) 100 *Journal of Political Economy* 5, 992–1026. Reputational cascades are different, as they concern someone's reluctance to express a certain judgment out of fear of socially falling from grace, cf. Timur Kuran and Cass R. Sunstein, "Availability Cascades and Risk Regulation", in: (1999) 51 *Stanford Law Review* 4, 683–768.

[328] Spiekermann and Goodin, "Courts of Many Minds", op. cit. 244, 558.

[329] Ibid.; Vermeule, *Law and the Limits of Reason*, op. cit. 326, 5-6.

Spiekermann and Goodin produced a formal computational model on courts and cascades in order to clarify the problem with the many-minds argument based on sequential judgements.[330] As regards traditionalism (and cosmopolitanism as much as it is traditionalist), their paper differs between homogenous courts, in which all judges observe the votes of their predecessors and also reveal their private signals, and heterogenous courts, in which some judges reveal their private signals and others aggregate the prior votes, but the latter must know whether the prior judges followed their own informational signal or the majority of past decisions, i.e., another aggregated (traditionalist) vote.[331] For the latter assumption to work empirically, judges need to be able to read the opinions of previous judges, which Spiekermann and Goodin hold to be easily the case in their American context.[332] This is more problematic in the European context, for example, due to the CJEU's practice of not issuing dissents or concurring opinions. For homogenous courts, the many-minds argument has to be treated with care when the voting is sequential as it is in traditionalism:

> "If judges are quite responsive to the opinions of their predecessors, they can quickly trigger cascades, compromising the capacity for many minds to enhance group competence. If they are more likely to reveal their own private signal in their decision because they are less responsive to previous judgements, they are not using the previous decisions to improve their own vote, but later judgements can benefit because early cascades are prevented and the information from more independent assessors is taken into account."[333]

In order to achieve the epistemic power of the many-minds, judges in homogenous courts would thus have to largely resist tradition.[334]

---

[330] In their model, judges are not fully Bayesian rational but merely boundedly rational in order to align their model with empirical findings of judges' behavior as being more likely to hold on to their own view than to defer to majorities, cf. Spiekermann and Goodin, "Courts of Many Minds", op. cit. 244, 558-559; Jeffrey A. Segal and Harold J. Spaeth, "The Influence of Stare Decisis on the Votes of United States Supreme Court Justices", in: (1996) 40 *American Journal of Political Science* 4, 971–1003.

[331] Spiekermann and Goodin, "Courts of Many Minds", op. cit. 244, 563.

[332] Ibid.

[333] Ibid., 562.

[334] Ibid.

For heterogenous courts, their results are not without irony: They can improve their chances of reaching correct decisions by their judges being traditionalists, but they benefit epistemically only if the votes of previous judges their judges are taking account of were not themselves traditionalist and the previous judges voted purely on basis of their own private signal.[335]

The authors argue that their results could be interpreted as either providing a starting point for an empirical analysis of the functioning of courts – to what extend to judges consider previous judgments and do they distinguish between traditionalist votes and non-traditionalist ones? – or as the starting point of a normative argument in favor of diversity in courts. According to their models, the epistemic performance of a court is poor if judges are traditionalists but also fairly poor if they do not consider precedent at all. Mixing interpretational approaches delivers better results.[336]

Overall, and including populism which we bracketed for our purposes here, Spiekermann and Goodin come to opposite conclusions as Sunstein does, which is remarkable given that both base their findings on the mechanisms of the CJT. From the perspective of truth-tracking, Spiekermann and Goodin find Sunstein's preference for traditionalism "questionable", as "one needs to provide arguments about how the problems arising from sequential voting can be avoided."[337]

Finally coming back to our question as to whether or not we can draft an Epistemic Principle of Judicial Deference based on the mechanisms of collective wisdom, the answer must be a mixed one. Let us therefore assume for now that ICs are heterogenous. An Epistemic Principle of Judicial Deference would then only suggest itself if a traditionalist court deferred to a non-traditionalist court.

Unless a previous judgment touches on the very same factual question, which should be rare for ICs, cases under empirical uncertainty are largely irrelevant for epistemic judicial deference. Normative uncertainty, however, might prove to be different. Let us thus consider the *Google Spain* judgment by the CJEU. Could a more reliable proportionality test have been available if the CJEU had drawn on the reasoning of the

[335] Ibid., 564.
[336] Ibid.
[337] Ibid.

ECtHR. After all, Article 7 of the Charter replicates Article 8 of the ECHR.[338] Moreover, due to the relative recency with which the CJEU gained jurisdiction over the Charter in 2009, compared to the ECtHR the CJEU had "little experience" in adjudicating fundamental rights in those early years.[339]

Yet from the judgment it is not recognizable if, or to what degree, the jurisprudence of the ECtHR played a role in the CJEU's reasoning, as the letter did not refer to the former (or the ECHR for that matter). For legal scholars, this is hardly surprising. In a study by Gráinne de Búrca, the CJEU demonstrated a "remarkable lack of reference on the part of the Court of Justice to other relevant sources of law and jurisprudence"[340], including the reasoning of the ECtHR, in the first three years after the entry into force of the Charter. We do not know, however, if and how much the CJEU's *référendaires*, judges and the Advocates General implicitly relied on the prior judgments of the ECtHR.

As argued above, a chamber-size comparison between the CJEU and the ECtHR is unrewarding. Likewise, a comparison based on cognitive diversity may see the ECtHR slightly ahead, but the question remains whether this is of any meaningful magnitude. The ECtHR can draw on judges from 47 different countries as opposed to 28 countries for the CJEU. However, the value diversity amongst the Council of Europe may prove to be much higher than in the EU, thus frustrating some of the benefits of cognitive diversity. As argued above, our indicators 'greater numbers' and 'greater diversity' prove to be not fine-grained enough to make an inter-court comparison as to both courts' epistemic reliability. This is not necessarily the case because of the way the mechanisms of collective wisdom work. Rather, it is the outcome of our categorical understanding of cognitive

---

[338] Cf. Shazia Choudhry, "Article 7 (Family Life Aspects)", in: Steve Peers, Tamara Hervey, Jeff Kenner, and Angela Ward (eds.), *The EU Charter of Fundamental Rights: A Commentary* (Oxford: Hart, 2014), 183–221, 201.

[339] Cf. Gráinne de Búrca, "After the EU Charter of Fundamental Rights: The Court of Justice as a Human Rights Adjudicator", in: (2013) 20 *Maastricht Journal of European and Comparative Law* 2, 168–184, , 170. This is not to say that the CJEU did not have any fundamental rights jurisprudence before the Charter, just see Case C-11/70, *Internationale Handelsgesellschaft mbH v. Einfuhr- und Vorratsstelle für Getreide und Futtermittel* [1970] ECR 01125. Some have speculated that the EU's choice to enact an EU Charter was motivated by preserving its own autonomy and exclusive authority, which might have been limited if a court outside of the EU interpreted binding fundamental rights, cf. de Búrca, "After the EU Charter", 172; .; see further: Opinion 2/94 [1996] ECR I-01759; Tobias Lock, "Walking on a Tightrope: The Draft ECHR Accession Agreement and the Autonomy of the EU Legal Order", in: (2011) 48 *Common Market Law Review* 4, 1025–1054; Christina Eckes, "EU Accession to the ECHR: Between Autonomy and Adaptation", in: (2013) 76 *Modern Law Review* 2, 254–285.

[340] De Búrca, "After the EU Charter", op. cit. 339, 173.

diversity. In this paper, the assumption is that a group comprised of two lawyers is less cognitively diverse than a group formed by a lawyer and an economist. Obviously, this does not have to be the case, depending on the people involved. We do need empirical evidence as to whether economists actually solve problems differently when sitting on a judicial bench.

Given our assumptions, however, an Epistemic Principle of Judicial Deference could only be endorsed if it was applied to the precedent of a non-traditionalist court. This, however, is at least questionable with a view to the normative reasoning of the ECtHR.[341]

d. An Epistemic Principle of Technocratic Deference

Technocratic institutions of public authority are, like courts, expert institutions. Naturally, their expertise is far removed from the legal expertise of judges. The European Commission, for example, is staffed with plenty of trained economists. We do not need to draw on the mechanisms of collective wisdom to establish that those technocratic institutions are generally epistemically more reliable in their domain than courts (whether or not the same holds true for parliaments is not the issue here). Problems of institutional legitimacy arise, however, when the competence of a technocratic institution is questioned in court. Is judicial epistemic deference justified even in those special cases?

The answer cannot be comprised of epistemic reasons alone. Procedural considerations must come into play, as the epistemic point of view would quickly render the review of contested technocratic decisions elusive. Sufficiently large supermajorities or rules of unanimity might help avoid situations in which courts defer to technocratic bodies despite significant disagreement among the bench.[342] Again, if disagreement led to the illegitimacy of technocratic deference, then this does not follow from the epistemic grounds of collective wisdom. This does not, however, speak of its irrelevance.

---

[341] See further: Yonatan Lupu and Erik Voeten, "Precedent in International Courts: A Network Analysis of Case Citations by the European Court of Human Rights", in: (2012) 42 *British Journal of Political Science* 2, 413–439.

[342] We might draw some inspiration from an analysis of judgment aggregation in coauthoring scientific papers, which, after all, might share some resemblance to written judicial judgments as documents of collaborative reporting, cf Liam Kofi Bright, Haixin Dang, and Remco Heesen, "A Role for Judgment Aggregation in Coauthoring Scientific Papers", in: (2018) 83 *Erkenntnis* 2, 231–252.

Take the *Microsoft*[343] judgment of the General Court of the CJEU as an example. A Grand Chamber of (then) 13 Judges upheld the European Commission's decision according to which Microsoft had violated EU competition law for, first, refusing to supply interoperability information to competing suppliers of workgroup servers and, two, for illegally tying Windows Media Player to the Windows PC operating system.[344] In the aftermath of the decision, then-president of the General Court of the CJEU, Bo Vesterdorf, who presided over *Microsoft*, gave a lecture in which it became clear that there must have been substantial disagreement on the bench.[345] Vesterdorf criticized that the GC had expanded the scope of conditions which had been previously established in the competition case law in order to identify the abuse of dominance Microsoft was found to have committed.[346] Steering away from the intricate details of competition law, it suffices to note that the expansion in scope included economic assessments which are generally within the broad margin of appreciation the CJEU grants the Commission in complex economic matters.[347]

The CJEU does not issue dissenting opinions and thus hides the disagreement which according to Vesterdorf must have been present in *Microsoft*. From the outside, however, we cannot tell how far-reaching the disagreement was. If Vesterdorf was the only Judge to disagree – and we do not even know whether he would have dissented or concurred – would we want a single judge among 13 Judges be enough to delegitimize the CJEU's decision? Would it be different if the *Microsoft* split had been 7:6? Given the vast discrepancy between the epistemic reliability in the domain of expertise at issue here, an Epistemic Principle of Judicial Deference would almost always point into the direction of the technocratic public authority. The example of the Epistemic Principle of Technocratic Restraint thus shows the importance of procedural to limit the hollowing-out of judicial oversight of technocrats via purely epistemic principles of reasoning.

---

[343] CJEU, T-201/04, *Microsoft Corp. v. Commission* [2007] ECR II-03601.

[344] Ibid.; Cf. the summary in Zhang, "The Faceless Court", op. cit. 191, 117.

[345] Bo Vesterdorf, "Article 82 EC: Where Do We Stand After the Microsoft Judgment?", in: (2008) 1 *Global Antitrust Review* 1, 1–14; Zhang, "The Faceless Court", op. cit. 191, 118–120.

[346] Vesterdorf, "Article 82 EC", op. cit, 345, 7; Zhang, "The Faceless Court", op. cit. 191, 118–119.

[347] Cf. *Microsoft*, op. cit. 343, paras 87–88, reiterating the standards set out in: CJEU, Joined Cases 56/64 and 58/64 *Établissements Consten S.à.R.L. and Grundig-Verkaufs-GmbH v. Commission* [1966] ECR 00429; CJEU, Case C-194/99 *Thyssen Stahl AG v. Commission* [2003] ECR I-10821; CJEU, Case C-12/03 P, *Commission v. Tetra Laval* [2005] ECR I-00987, para 39.

e. An Epistemic Principle of Judicial Activism

Lastly, we will consider how ICs could legitimately be activist. Usually the term 'judicial activism' is used with a negative connotation as depicting some form of overreach by judges. To the contrary, an Epistemic Principle of Judicial Activism follows a neutral understanding of judicial activism, merely recognizing another option for the judiciary to decide the case at hand. In the literature, different definitions of judicial activism have been offered, each accentuating the particular criticism they convey. Following an identification effort by Keenan Kmiec,[348] the most interesting for the practice of ICs are three closely related definitions: The first one is that of judicial legislation, where the critics think that the decision should have been left to the legislature. The second definition is that of a court departing from an accepted canon of interpretation. Finally, the third definition understands activism as result-oriented judging such as, in the case of the CJEU, following its own vision for Europe in its decision making.

As much as the institutional dimension of legitimacy is concerned, courts can be activist if their epistemic reliability is at least as high as that of the institution which has practical authority over the domain at issue. In this regard, decision in *Google Spain* seems hardly legitimate from an epistemic point of view if thought of as judicial legislation. The complex balancing of interests does not put the judges in an *a priori* better position than the European legislature. At least ehen compared with parliamentary decision making, the Epistemic Principle of Judicial Activism appears to be the inverse of the Epistemic Principle of Democratic Deference. While this seems to be correct in the institutional dimension of legitimacy – and thus presents an improvement when compared to von Staden's normative subsidiarity – epistemic reliability seems hardly to be the only factor for legitimizing judicial legislation. In this regard, normative subsidiarity in the procedural democratic sense of von Staden can still add value to the overall legitimacy assessment.

[348] Keenan D. Kmiec, "The Origin and Current Meanings of 'Judicial Activism'", in: (2004) 92 *California Law Review* 5, 1441–1478, 1471–1475. See also: Anthony Arnull, "Judicial Activism and the European Court of Justice: How Should Academics Respond?", in: Mark Dawson, Bruno De Witte, and Elise Muir, *Judicial Activism at the European Court of Justice* (Cheltenham: Edward Elgar, 2013), 211–232.

IV. Conclusion

This paper has outlined the institutional dimension of legitimacy as regards the authority of ICs. It has presented its own framework called Public Epistemic Authority, or PEA, which aims to provide an epistemic benchmark for institutional legitimacy by drawing on theories of collective wisdom. PEA is not chasing 'the Truth' but seeks to allocate decision-making authority and inform institutional design by utilizing the comparative cognitive potential of ICs. This move requires abandoning an individualistic paradigm of judicial competence in favor of a collective approach.

PEA's determination of epistemic reliability through mechanisms of collective wisdom is necessarily of an *a priori* nature. As Hong and Page explain, the idea that people receive independent signals that correlate with the truth drives the mathematical necessity of collective wisdom – and creates the gap between theory and reality.[349] Collective wisdom is therefore not a guaranteed, but a potential outcome, given that the right conditions hold.[350] As argued in this paper, the *a priori* comparative test of legitimacy needs to be complemented by an empirical behavioral test of legitimacy, accounting for cognitive and motivational biases as well as other systematic distortions of the independence of truth-tracking signals. This paper therefore ends with a carefully positive verdict as to the usefulness of the mechanisms of collective wisdom for scholars of IC legitimacy. They led to the formulation of several epistemic principles of judicial reasoning and suggestions for institutional design.

The CJEU has served as our main instance of an IC to be scrutinized under PEA. As regards its reasoning, the European judiciary fared well when considering its handling of empirical epistemic uncertainty in *Afton Chemical*, but was found lacking in its resolution of the normative epistemic uncertainty in *Google Spain*. As regards its institutional design, it could increase its institutional legitimacy by introducing non-lawyers to the bench or by the changing its the working language from French to English. An actual institutional redesign of the European judiciary would have to balance its epistemic

[349] Lu Hong and Scott E. Page, "Some Microfoundations of Collective Wisdom", in: Landemore and Elster, *Collective Wisdom*, op. cit. 12, 56–94, 57. As much as diversity is supposed to make the crowd wise, huge diversity amongst the signals people receive is quite possibly a "heroic assumption" in democracies, cf. Page, *The Difference*, op. cit. 145, 192.

[350] Hong and Page, "Some Microfoundations of Collective Wisdom", op. cit. 349, 57.

reliability with the demands for judicial accountability and independence.[351] The benchmark of epistemic reliability is not necessarily at odds with any of the two values. Its normativity, after all, needs to be deduced from the moral dimension of legitimacy.

This first exploration proceeded with several simplifications which further research will need to address. For example, the comparison of institutions based on the CJT seems to require a refined formal apparatus. Trying to increase cognitive diversity can introduce higher levels of value diversity to the decision-making group, with potentially adverse effects to its epistemic reliability. The equasion of cognitive diversity with diverse professional training and countries of origin has to be confirmed empirically. Lastly, theorists will have to balance epistemic reliability with procedural demands of legitimacy. In democratic theory, this debate is in full swing. It is about time that scholars of judicial authority join the conversation.

[351] Often, judicial performance is understood as a matter of judicial independence, cf. Chalmers, "Judicial Performance, Membership, and Design at the Court of Justice", op. cit. 159, 53. Cf., for example, Article 97 German Basic Law. Sometimes, it is even framed as a matter of judicial accountability and judicial authority is said to be exercised on behalf of the constituents, Chalmers, "Judicial Performance, Membership, and Design at the Court of Justice", op. cit. 159, 53. Cf. in this regard, CJEU, Case C-224/01, *Gerhard Köbler v. Republic of Austria* [2003] ECR I-10239.