The International Journal For Court Administration is an initiative of IACA’s Executive Board and its diverse membership. The Journal is an effective communications vehicle for the international exchange of experiences, ideas and information on court management, and contributes to improving the administration of justice in all countries. The collective international experience of its Executive Board and Editors has been that every judicial system, even in countries in the earlier stages of transition, has elements to it that may be of interest to others. The variations in practice and procedure from one region of the world to another, from one court system to another, also reveal major similarities across all systems. IJCA serves as a resource for justice system professionals interested in learning about new and innovative practices in court and justice system administration and management, in common law, continental, and Shari'ah-based legal systems throughout the world. The Editors publish two issues per year.

The Editors welcome submissions from court officials, judges, justice ministry officials, academics and others whose professional work and interests lie in the practical aspects of the effective administration of justice. To view the Editorial Policy and Procedures for Submission of Manuscript and Guidelines for Authors, visit the IACA website (www.iaca.ws) and chose IACA Journal.

The Journal accepts advertising from businesses, organizations, and others relating to court and justice systems by way of services, equipment, conferences, etc. For rates, standards, and formats, please contact one of the editors.

Riverboat court moored on Brazil’s Amazon River. Photograph by Dado Gladieri used with permission.
In this Issue:
From the Managing Editor:
Measuring Judicial Performance, Independence and Accountability
By Philip Langbroek

A Method for Assessment of the Independence and Accountability of the Judiciary
Frans van Dijk and Geoffrey Vos

Viewing Judicial Independence and Accountability through the “Lens” of Performance Measurement and Management
Ingo Keilitz

Conceptualization(s) of Judicial Independence and Judicial Accountability by the European Network of Councils for the Judiciary: Two Steps Forward, One Step Back
David Kosař and Samuel Spáč

Innovate – Don’t Imitate! - ENJC Research Should Focus on Research Gaps
Stefan Voigt

Researching judicial ethical codes, or: how to eat a mille-feuille?
Elaine Mak

Pitfalls in data gathering to assess judiciaries
Marco Fabri

Reaction on the comments on the ENJC study on Method for Assessment of Judicial Independence and Accountability.
Frans van Dijk and Philip Langbroek

From the Managing Editor:
Measuring Judicial Performance, Independence and Accountability
By Philip Langbroek, Managing Editor

On April 12-13, 2018, The European Network for Councils of the Judiciary, together with the Montaigne Center of Utrecht School of Law organized a seminar to discuss and validate a report on a major research project conducted by Frans van Dijk, Kees Sterk, Alain Lacabarats, and Guillaume Tusseau of the Netherlands and the French Councils for the Judiciary respectively, on the measurement of Judicial Independence and Accountability. The reports on this project can be found at the website of the ENCJ. This Special Issue is dedicated to that report and to that seminar, and guest editor Frans van Dijk and I acted as editors for this special Issue.

There are all kinds of methodological issues when trying to measure qualitative aspects of professional work. Measuring the Quality of Court Performance has been an exercise of the International Consortium for Court Excellence for about a decade. The Commission for the efficiency of justice (CEPEJ) of the Council of Europe, has been working for a decade on the measuring of the Quality of Judicial Systems. The Justice department of the European Commission publishes the Justice Scoreboard, with a set of indicators to enable an overview of the conditions of justice systems in EU –member states. And recently, a collection appeared on Measuring the quality of judicial reasoning, edited by Mátyás Bencze (Debrecen University, Hungary) and Gar Yein Ng (University of Buckingham), Springer, 2018. A year ago, a study appeared: Handle with Care, Assessing and designing methods for evaluation and development of the quality of justice, edited by Francesco Contini of the Research Institute of Judicial systems in Bologna, Italy.

Apparently, policymakers, court administrators and international institutions in the justice field, feel the need to show how they perform. In a world where politics has become budget centered, driven by financial policies, central banks and increasingly severe accounting policies, courts experience the increased focus on the results of their work, in their roles as part of the state, as provider of a public service, as wielder of state power and as countervailing power in relation to government and legislator. For scholars, developing adequate tools to measure judicial performances is a challenge and also part of an intellectual game they can take part in. For policymakers it seems also to be a tool to somehow get some control over the courts. Of course, as far as delivering justice by the courts is concerned, there is a very thin line between demanding improved service provision and pushing judgements in a specific direction. This may also be a threat in established rule of law democracies like the USA, France or the Netherlands. Organizing public accountability for courts and judicial performances in quantitative form seems to be the least harmful option from an independence and impartiality perspective. Representative democracies demand that also judiciaries account for how they perform, that is most reasonable. But independence and impartiality demand that political accounting mechanisms show respect for the constitutional positions of courts and judges, and their obligations to provide for fair trials. And although measuring performances within the court organisation may be helpful for court managers, it is in no way self-evident that the entirety of this information should be made public or shared with policymakers.

In the European civil law tradition, judges are civil servants, holding a professional public office. They fulfill their constitutional and trial function as a part of a profession, that takes as a point of departure that it should not matter who the judge in a case is. Judges are supposed not to participate in political debates and speak only through their judgements. In plural panel cases, a legal obligation to maintain the secrecy of chambers applies. The echo still resounds that judges are nothing than the mouth of the law. This makes European judiciaries vulnerable, and their impartiality and independence can only be realized adequately when the constitutional and societal functions of courts and judges are respected by the public, the media and furthered by politicians and court administrators. Of course, such a position comes with the
responsibility of courts and judges to provide the best possible services, living up to societal needs by timely quality work. Being open to criticisms by scholars and public debates is part of that responsibility. That is also why measuring court and judicial performances, and sharing the outcomes with the public and policymakers (at an aggregate level) makes sense. That enhances the public and political legitimacy of court- and judicial work.

This special issue shows some pioneering work in developing indicators of such indicators. This pioneering work comes with many remarks and criticism. Stefan Voigt, Ingo Keilitz, Marco Fabri, David Kosaf, Samuel Spáč and Elaine Mak react on and criticize the paper of Frans van Dijk and Geoffrey Vos: “A Method for Assessment of the Independence and Accountability of the Judiciary”. As Editors of this Special Issue, Frans van Dijk and I have commented upon the criticisms in the final contribution.

That is how our knowledge and experience grow.

Who is next in continuing those efforts?
A Method for Assessment of the Independence and Accountability of the Judiciary

Frans van Dijk and Geoffrey Vos

Abstract
The method outlined in this paper consists of a systematic assessment of the level of independence achieved in practice by national judicial systems throughout Europe, for the purpose of improving the design of judicial arrangements. Accountability of the judiciary is assessed alongside independence as those two concepts are intricately linked. Judicial independence and accountability are both evaluated in relation to the judicial system as a whole, as well as in relation to individual judges. Judicial independence is considered both objectively and subjectively. Specific indicators have been identified for each of these assessments. The whole methodology was applied to the judicial systems of 23 countries in Europe. The main findings of this exercise are summarised in this paper, with the strengths and weaknesses of these judicial systems and the most acute risks to judicial independence identified below.

Key words
Judicial Independence and Accountability, Measurement, Indicators, European Network for Councils of the Judiciary

TABLE OF CONTENTS
I. INTRODUCTION
II. PROBLEM DEFINITION AND METHODOLOGY
III. DEFINING AND MEASURING INDEPENDENCE AND ACCOUNTABILITY
   1. Judicial independence in relation to the other powers of the state
   2. Requirements of Judicial Independence
   3. Empirical data on Judicial Independence
   4. Independence and Accountability
IV. INDICATORS OF INDEPENDENCE AND ACCOUNTABILITY
   1. Framework
   2. Objective Independence of the Judiciary as a whole
   3. Objective Independence of the Individual Judge
   4. Subjective Independence of Judiciary and Judge
   5. Objective Accountability of the Judiciary as a whole
   6. Objective Accountability of the Individual Judge
   7. Methodology for the Measurement of the Indicators
V. OUTCOME OF APPLICATION OF THE INDICATORS
VI. USE OF THE OUTCOME OF THE INDICATORS
   1. Use at National Level: Example of the Netherlands
   2. Use at European Level: the Risks to Judicial Independence
VII. DISCUSSION

1 Frans van Dijk: Netherlands Council for the Judiciary and coordinator of the project teams about independence and accountability of the European Network of Councils for the Judiciary, and, since February 2018 professor at University of Utrecht, the Netherlands. e-mail: f.vandijk@uu.nl. Geoffrey Vos: Chancellor of the High Court of England and Wales, and President of the European Network of Councils for the Judiciary from January 2015 until June 2016. We are grateful to the participants of the ENCJ project teams Independence and Accountability, Ymkje Lugten, Monique van der Goes, Bart Diephuis, Frank van Tulder and Vannina Ettori for their contributions and the participants of the EGPA conference 2016 for their comments.
I. INTRODUCTION

Judicial independence is key to a properly functioning judicial system. It allows judges to make impartial decisions in accordance with law and evidence only, shielding them from inappropriate outside influence, whether from other branches of Government, the public, or the private sector. Independent judges are expected to be incorruptible and fearless; they should be able, where necessary, to decide cases in ways that may upset governments, media and public opinion. The impartial adjudication of disputes is a crucial component of the rule of law, and is essential to a peaceful, prosperous and democratic society. An independent judiciary promotes that citizens do not feel the need to take justice into their own hands, and that they can take short and long term financial decisions with confidence. This only works when the population actually observes and is convinced that justice is served in the way described. To make this possible the judiciary needs to function in a transparent and accountable manner, especially elucidating arrangements that impact directly independence, such as the manner in which cases are allocated to judges.

Judicial independence is entrenched in many legal instruments, such as the Universal Declaration of Human Rights and the European Convention on Human Rights and Fundamental Freedoms, and we take it here as starting point, including the connection between independence and accountability. Given the centrality of these concepts, it is important to have a clear understanding not only of what is required for judicial independence and accountability, but also to what extent these requirements are met in practice and to what extent judges actually behave independently. On the basis of this knowledge, the need for judicial reform can be assessed objectively, and informed decisions can be taken about the direction of judicial reform. This understanding feeds into the frequent dialogues that take place within the judiciary and between the judiciary and the other branches of government about the need of reform.

Opinion surveys by, in particular, the European Commission, World Justice Project and World Economic Forum show that judicial independence cannot be taken for granted. However, apart from such general surveys and the scientific literature discussed in this article, little has been done empirically to assess in depth the level of independence and accountability reached amongst the judicial systems of European Union member states, for the purpose of judicial reform. The European Commission's largely descriptive "Justice Scoreboard" has made a start in that direction while the Council of Europe has gathered data about some aspects of independence. In addition, a large literature exists that foremost aims at measuring judicial independence as a 'variable' to explain social/economic development. Central features are the distinction between 'de iure' and 'de facto' independence, and a search for measurable indicators of both (see further section III 3). While this literature contributes in many respect to the conceptualization and measurement of independence in general, it does not address the detailed design of judicial institutions. For the purpose of guiding judicial reform in practice, there is, therefore, a lack of systematic and comprehensive information about the diverse aspects of independence. It is for this reason that the European Network of Councils of the Judiciary ("ENCJ") set out in 2013 to make judicial independence, in combination with accountability, measurable, and in ensuing years developed its approach experimentally such that the results of this work can now be usefully presented for critical scrutiny and as contribution to the empirical analysis of judicial systems.

To be more precise, the ENCJ has developed a set of indicators based on European and international standards for judicial independence and accountability, using available data and collecting data otherwise. After a pilot application of this method in four sample states, the indicators were applied to 25 judiciaries in 2014-2015. A survey among the judges of Europe was part of the exercise. The results were evaluated and the indicators were improved and applied again in 2016-2017. For the ENCJ the indicators are primarily useful for improving judicial systems at the national level. Secondarily, the indicators provide insight into the overall situation of the judiciary in Europe. From this exercise, we have also been able to identify risks to the independence of the judiciary.

This article is structured as follows: Section II sets out the problem and the research questions that follow, and describes the methodology used to address these questions. Section III reviews the existing principles and standards about independence and the available data about independence to identify the elements needed for independence. It also examines the connection between independence and accountability, and the elements required for accountability. Section IV presents the indicators. It explains the content of the indicators and the method used for measurement. The results are presented in Section V, while in Section VI the application that can be made of these results at national level is demonstrated with the example of the Netherlands. The more general use of these results is also discussed in that Section, focusing on the risks to judicial independence. Section VII gives our conclusions on the effectiveness of this methodology.

2 UDHR art10 and ECHR art6.
3 The conceptualization of accountability is the subject of a large literature (see e.g. F. Contini and R. Mohr (2007), Reconciling independence and accountability in legal systems. Utrecht Law Review, http://www.utrechtlawreview.org, 3/2). As implied already, we use accountability in a broad sense to comprise - in the terminology of Contini and Mohr - foremost legal-judicial accountability and only secondarily managerial accountability.
6 For instance, data about the appointment of judges. Council of Europe CEPEJ, ‘European Judicial Systems - Efficiency and Quality of Justice’ (2016) CEPEJ Studies No. 23.
II. PROBLEM DEFINITION AND METHODOLOGY

To establish the need and direction of judicial reform in Europe a clear understanding of independence and accountability, theoretically and foremost practically, is necessary. This leads to three concrete questions.

1. What is required for a judge to be independent and accountable in a broad sense, personally and organizationally?

2. To what degree are the requirements met in the (candidate) member states of the EU, and are their judges actually independent and accountable? Do they behave accordingly?

3. What can be done to improve their independence and accountability?

As will be discussed in the next section, consensus exists at an abstract level about the answer to the first question among professionals and politicians in the international context of the UN, Council of Europe and international judges associations. It shows that independence and accountability are both multidimensional concepts that relate to how a wide range of matters are arranged in the law and applied in practice, and also how independence is perceived by (sections of) the population. Question 2 is another matter. Given the multi-dimensionality of both concepts, it requires the measurement of independence and accountability in a comprehensive manner. To what degree are the diverse requirements for independence and accountability met and does that actually lead judges to fulfil their duties independently and in an accountable manner? While some useful components are available that will be discussed later in the review of the literature, a comprehensive assessment nor a methodology how to do so are available. Emphasis is on this empirical question. The answer to question 3 follows from 2, in as far 2 provides sufficient insight in the underlying mechanisms.

The methodology followed can be summarized in five steps:

Step 1: identify the relevant aspects of independence and of accountability and their constituent parts on the basis of the existing international sets of principles and complemented by the actual experience of judges within the ENCJ.

Step 2: determine per aspect and component to what degree actual arrangements and practices are conducive to independence and accountability and rank them accordingly. Define the arrangements to be strived at.

Step 3: condense 1 and 2 into a set of indicators, consisting of (1) sub indicators that each can be categorized and (2) scoring rules to map the categories on quantitative scales. The indicators are about formal and factual arrangements on the one hand and about perceptions on the other hand. The formal and factual arrangements can be categorized and scored, at the level of sub indicators, directly and, in principle, unambiguously by any knowledgeable observer. The indicators that relate to perceptions of (aspects of) independence derive from opinion surveys, and are therefore quantitative by nature.

Step 4: gather all data for the countries of Europe that participate. As to the indicators about formal and factual arrangements the categorization was done by the Councils of the judiciary and, where Councils do not exist, by alternative governing bodies, such as ministries of Justice. As to the indicators about perceptions existing, primarily international surveys were used. In addition, a survey was held among the judges of Europe about their independence. The scoring of the indicators was done by the project staff, led by the Netherlands Council for the judiciary. The results are presented in the form of country profiles of the independence and accountability of the 25 judiciaries of 24 nations.8

Step 5: establish for which aspects outcomes for all participating judiciaries individually fall short of the aspiration levels, and identify potential remedies.

It should be noted that this approach attempts to generalize national experiences and therefore largely abstracts from historic, cultural and political background. This means that, before introducing reforms, decision makers in a nation have to assess whether short falls are actually problematic and remedies feasible under specific national conditions. Furthermore, one has to realize this is a first attempt at measuring independence and accountability in depth, and a critical examination of the country profiles is in order before embarking on reforms.

III. DEFINING AND MEASURING INDEPENDENCE AND ACCOUNTABILITY

1. Judicial independence in relation to the other powers of the state

The protection of judicial independence is guaranteed in core international instruments, in particular the Universal Declaration of Human Rights and the European Convention on Human Rights and Fundamental Freedoms. Both set out a person’s right to be heard by an independent and impartial tribunal in relation to criminal charges made against that person and/or the determination of that person’s rights and obligations.9 From this follows directly the independence of the judge, but the independence of the judiciary as a whole is not immediate.

---

8 For the UK the judiciary of England and Wales and that of Scotland were distinguished.
9 UDHR art10 and ECHR art6.
In this regard the position of the governments of Europe together is expressed by the Committee of Ministers of the Council of Europe in its Recommendation to Member States on Judges: Independence, Efficiency and Responsibilities. It states i.a. that “The independence of individual judges is safeguarded by the independence of the judiciary as a whole. As such it is a fundamental aspect of the rule of law.” This safeguard function is elaborated, for instance, in: “Where judges consider that their independence is threatened, they should be able to take recourse to a council for the judiciary or another independent authority, or they should have effective means of remedy.” The same is encapsulated in instruments and documents of the EC, in particular the EU Framework to strengthen the rule of law. The argument for independence of the judiciary is also made by judges. Lord Thomas (2017) argues the independence of the judiciary, while stressing the interdependence of the state powers. He sees as essential characteristics of interdependence a clear understanding by each state power of the constitutional functions and responsibilities of the other state powers, mutual support in carrying out these functions and no interference in each other’s functions.

While the above resolution does not elaborate its reasoning, it is only logical that, for the judiciary to play its constitutional role in the checks and balances between the state powers in a democracy, it should not be under the control of the other state powers. In the often invoked metaphor of Ulysses and the Sirens in constitutional theory: if Ulysses could order the untangling of the knot that binds him to the mast the whole idea of self-imposed self-restraint fails. If (new) governments or parliaments could intervene in the judiciary, either at the individual or organizational level, each time they did not like judicial decisions, the institutional arrangements are not fit for difficult times, and the balance of state powers fails. The result would be authoritarian rule with a subservient judiciary that cannot guarantee the fundamental rights mentioned at the start.

Interventions are generally at the organizational level such as reducing the pension age of judges and changing the selection/appointment procedures for Councils for the judiciary that are responsible for the selection, promotion and disciplining of judges. These examples are taken from recent events in Poland. Control over the judiciary can also be achieved by the other state powers through budgetary means, for instance, by underfunding the courts, as a result of which the courts cannot function properly. The duration of cases may become unacceptable for society or the courts cannot explain decisions to the public. In both situations public support for the judiciary is likely to decline, and this in turn facilitates further interventions.

The examples of appointments and budgets shows that there are always linkages between judiciary and the other state powers. Whether the independence of the judiciary is respected depends very much on the actual arrangements. The same holds with respect to organizational structure. The mere existence of a council of the judiciary is, for instance, no guarantee of independence. If the (judicial) members are selected by government or parliament, it is an extension of the other state powers within the judiciary. On the other hand if its members are elected by the judges, a council is a safeguard. Again, the actual arrangements are crucial. In the next section we explore these issues in detail by examining what exactly constitutes the independence of judge and judiciary.

2. Key Elements of Judicial Independence

What elements are required for judicial independence is the subject of a substantial set of documents, each more or less authoritative in nature. These documents exhibit considerable overlap, reflecting a broad communis opinio among governments and judges about what is needed for independence. Looking just at UN and European institutions related texts, the UN Basic Principles on the Independence of the Judiciary of 1985 provides a useful starting point. It establishes the principle that judicial independence must be legally enshrined and universally respected, that judges must be able to do their duty without pressure or improper influence, that there should not be revision of judicial decisions outside appeal and that judges should have immunity from civil liability. Another key issue identified in this document is the selection, promotion, removal from office and disciplining of judges. All these matters require extensive safeguards to be put in place to shield the judiciary from undue influence from, in particular, the other branches of government. Further issues relate to the terms and conditions on which judges operate. The resources allocated to the courts need to be adequate, and the security of office provided by tenure and similar mechanisms, as well as the level of remuneration afforded the judiciary need to be guaranteed in law. Although the UN document has been characterized as no more than the most basic framework, the principles capture most of the aspects included in other sets of principles.
Several international organizations of judges have also considered these matters. The European Association of Judges has developed a Judges’ Charter in Europe\textsuperscript{17} and the Consultative Council of European Judges\textsuperscript{18} a Magna Carta for Judges, while the International Association of Judges has a Universal Charter of the Judge\textsuperscript{19}. We focus here on what these documents add to the UN principles outlined above. The EAJ emphasizes that judges should not only be independent, but also be seen to be independent. That last element gives independence an extra dimension, as perceptions are of a different nature to formal, legal requirements.

The IAJ expands on internal independence: in their decisions, judges should be independent from other judges and from the judicial administration.\textsuperscript{20} The EAJ introduces governance and administration of the judiciary as another factor that impacts on judicial independence.

In its view, judicial administration should be carried out by a body which is representative of the judges and independent of any other authority. The CCJE addresses governance in a more specific manner: to ensure independence, each state should create a Council for the Judiciary or other specific body with broad competences that is independent from the legislative and executive branches of government.\textsuperscript{21} Another governance issue is raised: the judiciary should be involved in all decisions by government that affect the exercise of judicial functions such as the organization of the courts and the design of court procedures.

As mentioned above, the Committee of Ministers of the Council of Europe\textsuperscript{22} has made an important and authoritative statement about these matters. It stresses the importance of good governance for the judiciary and in particular the position and role of Councils for the judiciary. It also stresses the importance of internal independence: the existence of a hierarchical judicial structure should not undermine the independence of the individual judge. Its Recommendations set as principles that no cases be withdrawn from judges and that the judges themselves should be irremovable. Involuntary transfers to, for instance, other courts are unacceptable, with some exceptions, such as disciplinary measures and lawful re-organization. The Recommendation of the Committee of Ministers was informed by the Venice Commission of the same Council of Europe in a report that provides a summary of those principles. The Venice Commission addresses the need for adequate resources, as do most of the other documents, but it adds that the judiciary should be involved in the determination of its budget.\textsuperscript{23}

All these principles are generally formulated in an abstract manner, and in most cases do not specify what arrangements and mechanisms are needed for those principles to be met. Some aspects have been elaborated, for instance by the CCJE and the Venice Commission. See for instance the latter’s report on judicial appointment.\textsuperscript{24} Since Councils for the judiciary typically play an important role in putting the principles into practice, the European Network of Councils for the Judiciary has been working to develop those principles into standards and issue guidelines. An overview of its guidelines is given in ENCJ (2016)\textsuperscript{25}. Among other items, the selection, appointment and promotion of judges, their remuneration, the composition of Councils for the judiciary and court funding are addressed in the overview.

The above has implications for the design of a framework for the measurement of judicial independence. First of all, what is required for a judge to be independent is multi-dimensional, as there are many ways to improperly exert influence. It concerns the independence of the individual judge (e.g. human resource policies and internal judicial independence), but also of the judiciary as organization (e.g., governance and funding). In section IV About these dimensions broad consensus exists at the international level (UN, CoE, EU) between governments and judges organizations.

Secondly, the principles discussed above are largely derived from logic. For instance, judicial independence prevents state powers intervening in decisions related to the dismissal of judges. The dismissal of a judge for making a decision that goes against the interests of the executive would be directly contrary to the principle of judicial independence. Such principles are clear and can be strictly applied. Other principles reflect opinions about how the judicial system works or should work, based on, for instance, the shared views and experiences of judges. Matters surrounding the governance of the judiciary (Councils for the judiciary, other governance structures) fall into this category, and a framework should allow for situations where more than one solution exits.

Thirdly, the principles listed above consist mostly of formal requirements (i.e. how matters should be arranged in law), but also raise the issue of the perception of independence. Independence must be perceived by the population or, as suggested in the Bangalore Principles, by ‘a reasonable observer’, to exist.\textsuperscript{26} This allows, for instance, for the possibility that a legal system may

\textsuperscript{17} EAJ, see fn 16.
\textsuperscript{18} CCJE, Magna Carta of Judges (2010). This document summarizes the main conclusions of the Opinions that the CCJE has already adopted.
\textsuperscript{19} International Association of Judges, The Universal Charter of the Judge (1999).
\textsuperscript{20} See also the Bangalore Principles about Judicial Conduct (1.3): ‘A judge shall not only be free from inappropriate connections with, and influence by, the executive and legislative branches of government, but must also appear to a reasonable observer to be free thereof.’ As endorsed by UN ECOSOC 2006/23.
\textsuperscript{21} See also for a more elaborate view on judicial administration: OSCE ODIHR and Max Planck Minerva Research Group on Judicial Independence, ‘Kyiv Recommendations on Judicial Independence in Eastern Europe, South Caucasus and Central Asia: Judicial Administration, Selection and Accountability’ (2010).
\textsuperscript{26} Bangalore Principles about Judicial Conduct. See fn 20. Art 1.3.
possess all the formal requirements of independence and put these into practice, but nonetheless have judges who do not act independently. A framework needs to take this into account. In the literature, a related distinction is made between de jure and de facto independence, where perceptions may be a proxy for de facto independence (see the next section).

3. Empirical Data on Judicial Independence

Defining the elements that make up judicial independence is one thing, but assessing their actual implementation is another matter. Relatively little has been done in this area. The Council of Europe through CEPEJ provides data about some important matters such as the mechanisms and procedures for the selection and appointment of judges, terms of office (tenure and salaries) and the rules regarding the irremovability of judges. The European Commission issues an annual EU Justice Scoreboard. It pays attention to several aspects of independence, mentioned above. Data on the independence of courts and judges, as perceived by the general public and by companies, is presented. Sources are the Eurobarometer and the Global Competitiveness report of the WEF (see below). Descriptive information about the rules guiding the appointment of judges, the individual evaluation of judges, the transfer of judges without their consent and the dismissal of judges is provided. Also, some related quantitative data are included, for instance about dismissals. However, a systematic consideration and measurement of the matters discussed above does not exist.

Given this lack of systematic information, it is useful to examine different approaches for measuring independence. There is a body of literature that considers how to measure formal (de jure) independence and the rather elusive actual (de facto) independence of judges and the relationship between these measurements. This literature views judicial independence as a variable to be measured for use in, especially, models of economic development. This is, of course, very different from our concerns. Rios-Figueroa and Staton summarize the literature in this field. The indicators vary wildly and often concern very specific aspects. Nevertheless, they can provide insight into the wider ramifications of independence. ‘De jure independence’ is usually measured by indicators of some of the elements discussed above. Fixed tenure, objective appointment procedures, budgetary autonomy, and judicial councils are used for this purpose.

De facto independence is conceptualised as judges not responding to undue pressure to resolve cases in a particular way. A second interpretation of de facto independence involves judges’ decisions being enforced even when political actors would rather not comply with those decisions. Neither idea is easy to measure. Direct observation and analysis of decisions are cumbersome and may be misleading. Decision-making may appear autonomous when all the while there is some case selection at play and controversial cases are being removed from the court lists. Alternative approaches to measuring de facto independence have been developed. The first measures the conduct of actors in society as a reflection of the de facto independence of the judiciary. Indicators generally relate to actual constraints on executive authorities. In a very specialized variant of this approach, Clague et al use the ratio of non-currency money to the total supply of money as a measure of the trust within society in judicial institutions that enforce contractual obligations set by the banking industry. The second approach is based on perceptions of judicial independence as a proxy for de facto independence. These perceptions are derived from questionnaires circulated among respondents that have experience with the courts, such as lawyers and company executives. The World Economic Forum data mentioned above are based on this approach. Another source is public opinion data, e.g. the European Barometers.

A third approach to measure de facto independence focuses on indicators based on ‘objective’ information about the actual processes in place within the judiciary. This information is generally gathered by experts. Hayo and Voigt consider primarily the position of judges in the highest court (the court of last resort) in a country. Indicators are the average length of tenure of these judges, deviations from legal rules in this area, changes in the number of the judges, changes to the legal foundations establishing the highest court, and the degree of implementation of decisions of these courts. Additional factors include the growth in the income of judges and court budgets. However, the relevance of these indicators for actual judicial independence remains open.

---

30 J Rios-Figueroa, JK Staton. See fn 7.
32 Rios-Figueroa and Staton, see fn 7.
for discussion. As we saw in Section II, judicial independence is a multi-dimensional phenomenon, and it is questionable whether independence can be captured by a few limited variables.

An important issue addressed in the literature is the link between (indicators of) de jure and de facto independence. Hayo and Voigt (2005) find that their measures of de facto independence are partly dependent on ‘de jure independence’, but also on factors such as the confidence of the public in the judiciary, the extent of democratisation, the degree of press freedom and other cultural factors, like the religious beliefs of the population. Melton and Ginsburg (2014) find that rules governing the selection and removal of judges are the only de jure protections that actually enhance judicial independence in practice. Other studies find that de facto independence and de jure independence are quite different phenomena. Ríos-Figueroa and Staton conclude that Indicators of de jure and de facto independence are at best weakly correlated and that in some cases, these are even negatively related.38 They also find that different de jure indicators are not strongly correlated with each other. The correlations between the various indicators of de facto independence are, according to these authors, reasonably strong, despite their differences in content. Also studies of specific countries indicate that both concepts of independence have weak links with one another. In Romania de jure independence could increase without substantial effects on de facto independence.39 In Venezuela it was the other way round.40

We conclude that, while the concepts of de jure and de facto independence have received recognition, the conceptualisation and measurement of both are complicated. The indicators for de jure independence coincide with some of the aspects of independence discussed in the previous Section, but not comprehensively. While capturing important aspects, they do not do full justice to the multi-dimensionality of independence. De facto independence should essentially show in the content of the judgments, but that is extremely hard to grasp, and most studies resort to indirect measures. For the moment, we consider the perceptions of independence of different groups in society as relevant in itself and as approximation of de facto independence.

4. Independence and Accountability

From several perspectives, there are strong arguments against looking at independence in isolation. Here we focus on the link with accountability. Accountability is used here in the sense of the judiciary being morally obliged to inform society about all aspects of its functioning and explain its policies, procedures and decisions. Accountability and transparency are then closely linked, and we use these terms interchangeably. While it is sometimes suggested that the judiciary can best protect its independence by not being transparent, as transparency opens the door to interventions of the other branches of government and forces a managerial, bureaucratic and efficiency oriented approach on the courts, there are arguments for the opposite.41 The positive link between independence and accountability has an ideological value-driven dimension (judges are public servants who fulfil their duties under public scrutiny), but stems also from the functioning of democracies. From a political science perspective, laws and their proper implementation are the result of functioning political processes.42 This also holds for the constitution and other entrenched laws that set out the arrangements for the organization and functioning of the judiciary and thereby establish formal independence. While political processes are outside the control of the judiciary, the outcome of these processes is affected by the way judges and judiciary function. It seems reasonable to suggest that, while many factors play a role, governmental and political support for the independence of the judiciary - reinforced by popular support - will generally be stronger the better the judiciary is seen to fulfill its duties, despite the inevitable errors that will occur occasionally. The public can only be sure that judges are independent and impartial if the judiciary functions in a transparent manner and proves its willingness to be accountable to society. Accountability can, therefore, be seen as a necessary condition for independence, at least in the longer term. The internal argument for accountability is related: systems that are accountable to no one are likely to have weaker incentives to improve themselves than systems that open up to the outside world and tackle criticism seriously. By emphasizing and strengthening independence solely judiciaries risk insulating themselves from society and becoming irresponsible to justified demands of society.

In the view of the ENCJ a judiciary that is not accountable to society will not be trusted by society and will thereby endanger its independence: ‘Independence must be earned. It is, by no means, automatic. The best safeguard is excellent and transparent

38 Ríos-Figueroa and Staton. Fn 7.
41 For an overview see Contini and Mohr (fn 3). Especially, the new public management approach often receives strong criticism from judges. For a balanced and empirical analysis see A. Lienhard and D. Kettinger, eds (2016), The Judiciary between Management and the Rule of Law; Results of the Research Project ‘Basic Research into Court Management in Switzerland’. Part of the project was a survey among participants in the Judiciary. While congruence of the expectations about what constitutes a good judiciary of judges, management and other groups in the courts was found, the dominant culture is that of the judges expecting the judiciary to be impartial, non-arbitrary, independent, incorruptible and trustworthy. Court management puts more weight on a judiciary that is integrated in society, customer-friendly, accessible, impact oriented and well-equipped.
42 See for a similar discussion about independence and accountability of regulatory agencies: Hans Bredow Institute for Media Research, Interdisciplinary Centre for Law & ICT (ICRI), Katholieke Universiteit Leuven, Center for Media and Communication Studies (CMCS), Central European University, Cullen International Perspective Associates (2012), ‘INDIREG FINAL REPORT; Indicators for independence and efficient functioning of audiovisual media services regulatory bodies for the purpose of enforcing the rules in the AVMS Directive.’
43 E.g., T. Koopmans (2003), Courts and political institutions; a comparative view. Cambridge University Press.
performance."^{44} Recent events in Poland underline that legal safeguards do not always guarantee independence, as laws can be changed, sometimes surprisingly quickly."^{45}

On this basis, it is logical that the indicators should not only consider independence but also accountability. Consequently, the requirements for the judiciary to be accountable need to be identified. Transparency and accountability have, however, received much less attention than independence. Still, the sets of principles reviewed in Section II provide relevant context, albeit haphazard. The Recommendation of Ministers of the Council of Europe stresses the duty of judges to give clear reasons for their judgments in language which is clear and comprehensible."^{46} It also states that ethical principles of professional conduct should be laid down in codes of judicial ethics that inspire public confidence in judges and the judiciary. Further, the Recommendation identifies the circumstances in which a judge is to withdraw from a case. This topic is considered in greater depth in the Bangalore Principles of Judicial Conduct that list these conditions, but also address acceptable accessional functions and seemingly conflicting financial interests judges may have without compromising their independence."^{47} The Magna Carta of Judges of the CCJE suggests that transparency requires that information be published on the operation of the judicial system."^{48} Also, the CCJE argues that it should be made clear what corresponds to judicial misconduct. We conclude that in this area accountability implies that the judiciary set clear rules accessible to the public to avoid its impartiality being compromised. This requirement does not only apply to individual behaviour. The Venice Commission pays much attention to case allocation that in its view should not be,"^{49} but based on objective and transparent criteria."^{50} Case allocation can be seen as a matter of internal independence, but it is also important from the perspective of transparency. Again, there need to be rules and these must be made public. The UN's basic principles refer to the need for a complaints procedure. While not comprehensive, these recommendations give an indication of the elements that need to be incorporated in a framework that not only aims to guarantee independence but also ensures judicial accountability."^{51} This short discussion should also make clear that accountability goes well beyond managerial accountability, focused on production and efficiency.

As to perceptions or other empirical data about accountability, not much is available."^{52} We are not aware of data that measures the perceptions in society about the accountability of the judiciary.

**IV. INDICATORS OF JUDICIAL INDEPENDENCE AND ACCOUNTABILITY**

1. **Framework**

Using the findings of Section III, a conceptual framework for the operationalization and measurement of independence and accountability has been developed. Table 1 lists the key aspects that have been identified. As discussed above, an important distinction is between formal or objective independence and subjective or perceived independence. Formal legal safeguards may make the judiciary objectively independent, but the subjective perception of judicial independence as viewed by different sections of society, including judges themselves, is equally important. The other major distinction is between the independence of the judge and that of the judiciary. The judge needs to be independent to decide cases impartially, but that also requires that the judiciary as a whole is independent, taking its interdependence with the other state powers into account. For instance, while parliament has budget right, insufficient and arbitrary funding of the judiciary can make individual independence an empty shell. Independence should therefore be defined at two levels. This distinction applies only to formal independence, as perceptions generally do not differentiate between the individual judge and the judiciary.

These two distinctions apply to accountability as well. There is a lack of data about subjective accountability, and this aspect has had for the moment to be excluded. The distinction between the individual judge and the judiciary is necessary. Systemic accountability requires transparency of procedure (e.g. case allocation and complaint procedures) and transparency of performance (e.g. timeliness and efficiency). Individual judges are accountable if their decisions are publicly available, properly reported and explained.

An indicator was devised for each key element of objective and subjective independence and objective accountability. Each indicator captures a complex phenomenon and consequently consists of several sub indicators. The next 5 paragraphs explain briefly each of the indicators. It needs to be stressed that the indicators are based on the common ground in published legal instruments and analyses of independence and accountability that we discussed in Section III and they reflect the consensus within the ENCJ in which nearly all judiciaries of Europe are represented. The indicators follow directly from the literature discussed in Section III.

---

46 CoE, Fn 22, Chapters VII and VIII.
47 UN/ECOSOC, Bangalore Principles of Judicial Conduct. Fn 20, Sections 2.5, 4.7 and 4.11, in particular.
48 CCJE, Fn 18.
50 For a similar combination of independence and accountability see also OSCE, Kyiv Recommendations. See fn 21.
51 Voigt arrives at the same conclusion, but chooses a different conceptualization of accountability then we do here. S. Voigt, The economic effects of judicial accountability: cross-country evidence (2008) European Journal of Law and Economics 25, 95-123,
Table 1. Overview of key aspects of judicial independence and accountability

<table>
<thead>
<tr>
<th>Formal independence</th>
<th>Perceived independence</th>
<th>Formal accountability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Judiciary:</strong></td>
<td><strong>Judge and judiciary undifferentiated:</strong></td>
<td><strong>Judiciary:</strong></td>
</tr>
<tr>
<td>· Legal basis of independence</td>
<td>· Independence as perceived by citizens</td>
<td>· Allocation of cases</td>
</tr>
<tr>
<td>· Organizational autonomy</td>
<td>· Independence as perceived by court users</td>
<td>· Complaints procedure</td>
</tr>
<tr>
<td>· Funding</td>
<td>· Independence as perceived by judges</td>
<td>· Periodic reporting by the judiciary</td>
</tr>
<tr>
<td>· Management of court system</td>
<td>· Trust in judiciary, relative to trust in other state powers by citizens in general</td>
<td>· Relations with the press</td>
</tr>
<tr>
<td><strong>Judge:</strong></td>
<td>· Judicial corruption as perceived by citizens in general</td>
<td>· External review</td>
</tr>
<tr>
<td>· Human resource decisions about judges</td>
<td></td>
<td><strong>Judge:</strong></td>
</tr>
<tr>
<td>· Disciplinary measures</td>
<td></td>
<td>· Code of Judicial ethics</td>
</tr>
<tr>
<td>· Non-transferability of judges</td>
<td></td>
<td>· Withdrawal and recusal</td>
</tr>
<tr>
<td>· Internal independence</td>
<td></td>
<td>· Admissibility of external functions and disclosure of external functions and financial interests</td>
</tr>
<tr>
<td></td>
<td></td>
<td>· Understandable procedures</td>
</tr>
</tbody>
</table>

2. Objective Independence of the Judiciary as a Whole

Applying the principles and their elaboration in guidelines discussed in section III, the key aspects of Table 1 have been developed into indicators, generally consisting of sub-indicators to capture the diverse elements that are relevant for a key aspect. The sets of principles about judicial independence overlap to such an extent that references for each indicator to specific sets of principles is irrelevant. The elaboration of the indicators is based on an analysis of the elements that are of practical importance in the judiciaries of Europe, utilizing the guidelines that have been drawn up by the ENCJ in particular (see fn 25; the guidelines are summarized in the Appendix of this article).

Table 2. Indicators of the objective independence of the judiciary as a whole

1. Legal basis of independence, with the following sub-indicators:
   - Formal guarantees of the independence of the Judiciary
   - Formal assurances that judges are bound only by the law
   - Formal methods for the determination of judges’ salaries
   - Formal mechanisms for the adjustment of judges’ salaries
   - Formal guarantees for involvement of judges in the development of legal and judicial reform

2. Organisational autonomy of the Judiciary, with the following sub-indicators where there is a Council for the Judiciary or equivalent independent body:
   - Formal position of the Council for the Judiciary
   - Compliance with ENCJ guidelines
   - Scope of responsibilities of the Council

   Sub-indicator when there is no Council for the Judiciary or an equivalent body:
   - Influence of judges on decisions in areas of responsibility

3. Funding of the Judiciary, with the following sub-indicators:
   - Budgetary arrangements
   - Procedure for resolution of conflicts about budgets
   - Objectivity of funding system
   - Sufficiency of actual budgets for categories of tasks
4. Management of the court system.
- Management responsibility of the courts for categories of tasks

There are the following indicators about the independence of the judiciary as a whole:

- the legal basis of judicial independence, which includes an enquiry about formal statutory guarantees of judicial independence and formal assurances that judges are bound only by the law. In addition, the indicator looks at the existence and application of formal processes for the determination and adjustment of judges’ salaries, and for the involvement of judges in government decisions about legal and judicial reforms that affect directly their tasks and functioning. In the absence of these guarantees the independence of the judiciary can be easily compromised and less readily enforced. Also, the nature of formal guarantees is relevant with constitutional guarantees providing the strongest protection.

- the arrangements that are in place to ensure the organizational autonomy and self-governance of the judiciary. The indicator examines the design and responsibilities of Councils for the Judiciary and of alternative institutional arrangements. 52 Apart from the way the formal position is arranged which determines the ease with which a Council can be abolished or altered, the indicator addresses whether arrangements are consistent with ENCJ guidelines about the set-up of an independent Council. 53 These guidelines elaborate the existing principles and concern for instance the composition of a Council. Among other: at least 50% of the members of the Council are judges and are chosen by their peers (see Appendix). This determines to a large extent whether a Council is part of the judiciary or an extension of government, and is a flashpoint in several countries (Poland, Spain). In addition, the indicator deals with the responsibilities of a Council. Even a perfectly appointed Council has little impact if its responsibilities are limited. The indicator allows for other institutional arrangements than Councils for the judiciary, in as far as these provide for decisive influence of the judges.

- the funding of the judiciary, including questions concerning the formal budgetary arrangements for both the judiciary as a whole and for court management, the processes in place to resolve budgetary conflicts, and the objective sufficiency of the actual budgets themselves to fund defined key tasks (see Appendix). As to the funding system, the use of objective criteria such as case load is seen as particularly important for independence, as this protects the Judiciary against arbitrary decisions by budget authorities.

- the method by which the court system is managed. In several countries the Minister of Justice is directly responsible for court management. The more that decisions in defined key areas are taken by the judiciary, the better independence is served, as will be discussed further in section 7.

3. Indicators of Objective Independence of the Individual Judge

Table 3. Indicators of the objective independence of the individual judge

5. Human resource decisions about judges, with the following sub-indicators:
- Selection, appointment and dismissal of judges and court presidents: authority to decide
- Selection, appointment and dismissal of Supreme Courts judges and the President of the Supreme Court: authority to decide
- Compliance with ENCJ guidelines about the appointment of judges
- Evaluation, promotion, disciplinary measures and training of judges: authority to decide
- Compliance with ENCJ guidelines about the promotion of judges

6. Disciplinary measures, with the following sub-indicators:
- Compliance with ENCJ standards about disciplinary measures against judges
- Competent body to make decisions about disciplinary measures against judges

7. Non-transferability of judges, with the following sub-indicators:
- Formal guarantee of non-transferability of judges
- In the absence of a formal guarantee: arrangements for the transfer of judges without their consent

52 The ENCJ not only represents the perspective of Councils for the judiciary; its observers have equal status in its project teams.
8. Internal independence, with the following sub-indicators:
- Influence by higher ranked judges
- Use and status of guidelines
- Influence by the management of the courts

The indicators here include:

- the way in which human resources decisions are taken in relation to judges. Who decides, on what grounds and are safeguards in place? The questions relate to judicial selection, appointment and promotion procedures, the dismissal of judges, and to processes for the evaluation and training of judges. Compliance with ENCJ guidelines that define elementary standards of transparency and objectivity is a major element here.

- the way in which disciplinary measures are taken against judges. This extremely sensitive area requires safeguards in terms of procedures and governing bodies. Also, in this area the ENCJ has developed guidelines.

- the ability to transfer judges between courts without their consent for other than for disciplinary reasons. There should be formal safeguards in place to prevent judges being moved or removed from office, as this can be easily used to affect the outcome of cases.

- the influence that senior judges and court management can wield over judges generally. Obviously, attempts of court management to affect the outcome of cases can only be viewed as very negative, while pressure to decide cases in a timely manner may be warranted in some cases but is still problematic from the perspective of independence. The indicator also addresses the development of guidelines for matters such as uniformity, consistency, timeliness and efficiency by judges at the same level. The indicator seeks to identify such practices, as these are from the perspective of independence problematic in particular if guidelines are binding.

4. Indicators of Subjective Independence of the Judiciary and Individual Judges

The indicators of the two previous paragraphs are about the degree to which formal and factual requirements of judicial independence are met. The current paragraph is about perceptions of independence, and requires a different methodology, using surveys amongst relevant stakeholders.

Table 4. Indicators of the subjective independence of the judiciary and the individual judge

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Source(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9. Independence as perceived by society;</td>
<td>Flash Eurobarometer 435 2016 ‘Perceived independence of the national justice systems in the EU among the general public’ and Flash Eurobarometer 436 2016 ‘Perceived independence of the national justice systems in the EU among companies’</td>
</tr>
<tr>
<td></td>
<td>Global competitiveness report 2016-2017</td>
</tr>
<tr>
<td></td>
<td>World Justice Rule of Law Index 2016</td>
</tr>
<tr>
<td>10. Independence as perceived by courts users at all levels;</td>
<td>National surveys</td>
</tr>
<tr>
<td>11. Independence as perceived by judges;</td>
<td>ENCJ survey, question 13</td>
</tr>
<tr>
<td>12. Judicial corruption as perceived by citizens in general;</td>
<td>EU Anti-Corruption Report 2014</td>
</tr>
<tr>
<td>13. Trust in Judiciary, relative to trust in other state powers by citizens in general;</td>
<td>National surveys</td>
</tr>
</tbody>
</table>

The indicators address perceptions of judicial independence amongst the population and business sector in general, court users and judges. The first two indicators are separate because many citizens will not have recent or direct experience of the courts. The perceptions of judges about their (and their colleagues’) independence are relevant, as these perceptions provide direct insight into the workings of the formal mechanisms that should guarantee independence. A comparison of perceptions of judges with those of the population and court users is relevant, if only to check the consistency of perceptions. Perceptions amongst the population about judicial corruption are also included, as corruption is a fundamental breach of independence. As discussed in Section III, judges must be able to do (and must do) their duty without pressure and improper influence. Lastly, in-country comparisons between the
trust of citizens in the judiciary and their trust in the executive and the legislature give a different, but relevant perspective. While it should be recognized that trust is circumstantially linked to independence and it is a complex notion in itself, it is the only measure that is available in many countries about the relative performance of the judiciary within a country. The other indicators provide cross country data. These cross country comparisons will to some extent be affected by the perceptions about public institutions in general. The in-country comparison allows a more nuanced view.

There are, therefore, the following indicators under this heading:

• the perceptions of citizens, ideally breaking them down into different economic, status and ethnic groups, but for now undifferentiated: reliance was placed on the Eurobarometer surveys about independence among citizens and companies, the World Economic Forum’s Global Competitiveness Report and the World Rule of Law Index.

• court users’ perception of judicial independence. This requires national sources.

• judges’ own views of their independence. The ENCJ undertook an extensive survey of EU judges at the end of 2016.

• perceived judicial corruption, as reported by the Eurobarometer 2014. Since there are obvious difficulties in measuring the level of actual corruption, the perceptions in society are a good secondary indicator.

• the trust that citizens place in the judiciary compared with their trust in the other branches of government. This requires national sources.

5. Indicators of Objective Accountability of the Judiciary as a Whole

This and the next paragraph are about the formal and factual aspects of accountability. In section III we discussed that, while broad consensus exists about the need and requirements of independence, accountability has received much less attention. Managerial accountability about production and efficiency is even a source of controversy. Still, in paragraph 3 of that section we were able to distinguish aspects of accountability in a broad sense on the basis of existing sets of principles. It should be recognized, however, that the indicators presented in these paragraphs are more open for debate than those about independence.

<table>
<thead>
<tr>
<th>Table 5. Indicators of the objective accountability of the judiciary as a whole</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Allocation of cases, with the following sub-indicators:</strong></td>
</tr>
<tr>
<td>- Existence of a transparent mechanism for the allocation of cases</td>
</tr>
<tr>
<td>- Content of the mechanism for the allocation of cases</td>
</tr>
<tr>
<td><strong>2. Complaints procedure, with the following sub-indicators:</strong></td>
</tr>
<tr>
<td>- Existence of a complaints procedure</td>
</tr>
<tr>
<td>- External participation in the complaints procedure</td>
</tr>
<tr>
<td>- Scope of the complaints procedure</td>
</tr>
<tr>
<td>- Appeal against a decision on a complaint</td>
</tr>
<tr>
<td><strong>3. Periodic reporting by the Judiciary, with the following sub-indicators:</strong></td>
</tr>
<tr>
<td>- Availability of annual reports</td>
</tr>
<tr>
<td>- Publishing of the annual reports</td>
</tr>
<tr>
<td>- Scope of the annual reports</td>
</tr>
<tr>
<td>- Existence of periodic and public benchmarking of the courts</td>
</tr>
<tr>
<td><strong>4. Relations with the press, with the following sub-indicators:</strong></td>
</tr>
<tr>
<td>- Explanation of judicial decisions to the media</td>
</tr>
</tbody>
</table>

54 Trust is a broader and less precise concept than perceptions of independence. For instance, Jackson et al. define trust in the criminal justice system as follows: “To trust in the police and the criminal courts is to assume that criminal justice agencies and agents are willing and able to do what they are tasked to do (...). Spanning both intentions and abilities, trust is the belief that individuals working for criminal justice institutions have appropriate shared motivations and are able to fulfill their roles competently (...).” J Jackson, J. Kuha, M Hough, B Bradford, K Hohl, M Gerber, Trust and Legitimacy across Europe: A FIDUCIA Report on Comparative Public Attitudes towards Legal Authority (2013) Fiducia.

55 EC, Eurobarometers. See fn 36.

56 World Economic Forum, see fn 35.

57 World Justice Project, Rule of Law Index 2016 (2016).


5. External review, with the following sub-indicators:
   - Use of external review
   - Responsibility for external review

The indicators in this category are as follows:

- **Case allocation**, which should guarantee the impartial and expert treatment of every case. The allocation of cases to judges can be misused by, for instance court management, to affect the outcome of cases by giving cases to judges that can be influenced or of which the opinions are in line with desired outcomes. The indicator questions the existence, nature and transparency of the case allocation mechanism. In this area ENCJ-guidelines exist. It should be noted that case allocation is also relevant from the perspective of judicial independence.

- **The existence, scope and nature of complaints procedures for parties or their legal representatives in relation to matters such as the judges’ case handling methods and the behaviour of court staff.** The indicator also addresses external participation in complaints procedures and the number of complaints actually made.

- **Periodic reporting by the judiciary.** Reporting at system level provides insight in the performance of the judiciary as a whole and makes external scrutiny possible. Performance is about judicial matters as well as managerial matters such as the use of funds and timeliness. See the Appendix.

- **The relationship between the judiciary and the media.** The indicator questions the existence of a transparent dialogue with the media to explain judicial practices and decisions, and the judiciary’s participation in education so as to explain the role of the judicial decision-making in society.

- **The extent to which the judiciary is open to external review, including audit, quality and efficiency evaluations.** The indicator considers the types of external review undertaken and the uses made of them.

6. Indicators of Objective Accountability of the Individual Judge

<table>
<thead>
<tr>
<th>Table 6. Indicators of the objective accountability of the individual judge</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>6. Code of judicial ethics, with the following sub-indicators:</strong></td>
</tr>
<tr>
<td>- Existence of a code of judicial ethics</td>
</tr>
<tr>
<td>- Availability of training on judicial ethics</td>
</tr>
<tr>
<td>- Responsible body to provide judges with guidance or advice on ethical issues</td>
</tr>
</tbody>
</table>

| **7. Withdrawal and recusal, with the following sub-indicators:** |
| - Obligation of voluntary withdrawal |
| - Sanction on breach of an obligation to withdraw |
| - Procedure for request for recusal |

| **8. Admissibility of external functions and disclosure of external functions and financial interests, with the following sub-indicators:** |
| - Existence of policy on admissibility of external functions |
| - Authorisation of the exercise of accessory functions |
| - Existence of a (public) register of external functions of judges |
| - Existence of a (public) register of financial interests of judges |

| **9. Understandable proceedings, with the following sub-indicators:** |
| - Duty of judges to make proceedings intelligible to the parties |
| - Duty of judges to make proceedings intelligible to categories of court users such as children, youth, disabled people (physically/mentally), victims, those for whom the national language is not their mother tongue and self-represented litigants |
| - Training of judges |
There are the following indicators under this heading:

- the existence of a code of judicial ethics that lays down the standards of conduct that society and parties in particular can expect from judges, and for which judges can be held accountable. Such an ethical code may promote a better understanding of the role of the judge in society, and thereby promote public confidence. The indicator also covers training on judicial ethics, as the mere existence of a judicial code is not sufficient to affect the behaviour of all judges.

- the circumstances in which an individual judge will voluntarily withdraw from a case and recuse himself. These circumstances may relate to objective matters such as family or friendship ties or to subjective issues such as expressed opinions. The indicator also queries the consequences of a breach of an obligation to withdraw, when citizens can request recusal, which authority decides on recusal, and appeals against recusal decisions.

- the policies on judges undertaking external paid and unpaid offices and functions, the types of permitted activities and interests, and the availability of a public register of such matters and of the disclosure of financial interests of judges. Judiciaries follow different strategies. In some countries judges are not or in very limited cases allowed to fulfill functions outside the judiciary; in other countries judges have much more possibilities. In the latter situation a public register of functions is needed. Both strategies are acceptable, and the indicator allows for this.

- the intelligibility of judicial proceedings, querying the existence of a duty on individual judges to make proceedings understandable to the parties and, in particular, vulnerable parties, and the training of judges in accessibility generally.

Assessment of subjective indicators of judicial accountability requires the existence of opinion surveys, preferably amongst a wide range of relevant groups. Whilst survey data exists in relation to independence, it does not to our knowledge at the moment exist in relation to judicial accountability, and the lacuna is less easy to fill than that in relation to the perceptions of judges. Accordingly, this aspect will require further attention in due course.

7. Methodology of the Measurement of the Indicators

The indicators about objective independence and accountability, as described in the paragraphs 2, 3, 5 and 6 of this section, have been developed into a precise set of questions that were answered by the Councils for the judiciary and other governing bodies. The questionnaire is available on the ENCJ’s website. A points system, consisting of scoring rules, was used to evaluate the answers to the questions. Above, we indicated already for most indicators what is good and what is bad practice. The underlying principles that were applied are discussed here.

First, the key issue in general about formal safeguards is the ease with which such safeguards can be altered or removed. A safeguard embedded in a constitution offers more protection than one contained in normal legislation. Legislative safeguards are more effective than those contained in subordinate legislation, general jurisprudence or tradition. This principle lies at the heart of formal independence. All sets of principles and other documents discussed in Section III require protections in law and often in the constitution.

Secondly, judicial self-government is in principle desirable. Where - on the contrary - other branches of government have the authority to make decisions about the judiciary, decisions based on objective criteria are to be preferred to discretionary decisions.

Thirdly, in every area, transparent rules are to be preferred to ad hoc reactions to particular situations. This applies equally to independence (i.a. funding of the judiciary) and accountability (i.a. case allocation and complaint procedures). Judicial decisions and procedures should all be formalised, public and transparent. In addition, transparency requires active dissemination of information, rather than simply making information theoretically available.

Finally, in several areas, such as the appointment and promotion of judges but also the autonomy of the judiciary, guidelines have been developed by the ENCJ, taking into account opinions by inter alia the CCJE and the Venice Commission. Adherence to the guidelines is obviously to be preferred.

Applying these principles results in a points system by which all aspects of objective independence and accountability can be evaluated. Equal weight was given to all sub-indicators. The actual allocation of points to the possible answers to each question was developed by a common understanding within the ENCJ, which proved possible despite the differences between legal systems and cultures. The points system is given in the Appendix.

The scores for the indicators of objective independence and accountability were derived from the questionnaire, whilst those for subjective independence were based on international opinion surveys. Those surveys included the survey undertaken by

61 See Lienhard and Kettinger (fn 41) about the courts’ right of self-administration, as included in the federal constitution of Switzerland for the Federal Supreme Court.
62 As an example: the first item of the questionnaire asks whether and, if yes, how independence of the judiciary or the judge is formally guaranteed: (1) constitution or equivalent document where equivalence means that the position of the judiciary cannot be changed by simple majority, (2) law that can be changed by simple majority and (3) constitutional court. Option 1 earns 3 points, option 2 earns 2 points, option 3 earns 1 point and no formal protection earns 0 points.
the ENCJ itself about the independence of judges. International surveys were, however, not available about the opinions of court users and the trust of citizens in the judiciary relative to their trust in the other state powers. For these perceptions national sources, as reported by Councils for the judiciary and, where these do not exist, other governing bodies such as Ministries of Justice, had to be relied on.

The questionnaires were completed either by Councils for the judiciary or other governing bodies such as Ministries of Justice. The answers were scored by the project staff. Thus, while the indicators measured in this way are objective in the sense that they capture the formal arrangements of a legal system, the method is self-evaluation. In some situations, this may lead to self-serving bias. This is difficult to avoid, but a group of experts from within the ENCJ responded to queries about the interpretation of the questions, checked the logic and plausibility of the answers, and resolved ambiguities, ensuring as far as possible that the indicators were measured uniformly and correctly.

V. Outcome of the application of the indicators

We have now described the indicators that were applied and the method by which they were measured. The indicators are primarily intended to allow individual Councils for the Judiciary to improve themselves by assessing their institutional design from the perspective of independence and accountability and their performance in these respects against a standard. The outcomes are, therefore, presented in the form of country profiles, but we first examine the average results across all participating judiciaries (see Figure 1). In total 23 judiciaries participated in 2016/2017. The outcomes for each indicator are presented as a percentage of the maximum possible score that reflects the best possible arrangements. Statistics such as mean and standard deviation can be calculated for each sub indicator as well as indicator over all countries. It is not possible to do this across the indicators. The score per indicator is given in combination with the minimum and maximum scores achieved by any of the participating judiciaries. As to the availability of data, all indicators could be measured for nearly all countries, except for the independence of the judiciary as perceived by court users. Surveys among court users are unfortunately still quite rare. As a result, most judiciaries have a minimum score on this indicator. Given the crucial importance of court user feedback, the indicator was retained.

The average outcomes for the indicators show tendencies within the judiciaries in Europe. Given the differences between the judiciaries that we will discuss later, the average scores give only a rough indication. Nevertheless, some general conclusions can be drawn from the averages in combination with a global inspection of the country outcomes. In the first place, there is much room for improvement with respect of independence as well as accountability, judging from the difference between the actual scores and what are deemed best arrangements (100%). For most indicators at least one judiciary reaches this best level (see the position of the green dashes in Figure 1), showing that these best arrangements are achievable. On the other hand, very low minimum scores also occur (red dash), especially in the area of accountability. In the second place, with regard to objective independence (shown in light blue on the left side in the figures), the funding of the judiciary and court management score lowest by far. The funding of the judiciary is generally not well arranged, and judiciaries are dependent on discretionary budget decisions by the other branches of government. Court management is still often in the hands of Ministries of Justice. It has proven difficult to change financial and organisational arrangements.

64 Austria, Belgium, Bulgaria, Croatia, Denmark, Finland, France, Germany, Hungary, Italy, Latvia, Lithuania, the Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, UK England and Wales, UK Scotland.
65 As interval scales are used (per sub-indicator points can be earned on a scale with equal intervals: the distance between 1 and 2 is the same as between 2 and 3), percentages are allowed.
In the third place, the lowest mean scores concern subjective independence (shown in dark blue in the figures). As mentioned already, most judiciaries do not conduct court user surveys. As a result, the average score on indicator 12 is very low. The score on corruption is also low. The scores on the other subjective indicators are, however, at similar levels as the indicators about objective independence. These indicators concern judicial independence as perceived by citizens and independence as perceived by judges themselves. The correlation between both types of perceptions is high\(^6\), showing that the perceptions of judges of their actual independence are reasonably in agreement with those of citizens. Indicator 10 warrants specific attention, because it provides an in-country perspective. It concerns trust of citizens in the judiciary relative to trust in the other state powers. In nearly all countries the trust in the judiciary is higher than the trust in the other state powers (16 of the 18 countries for which data exists).

In the fourth place, with respect to accountability (shown on the right side in light blue), outcomes vary considerably among countries. For instance, about half of the countries score very low on periodic reporting, whilst the others score very high. More generally, external review and (disclosure of) external functions of judges get low scores, again with substantial country exceptions. The external nature of external review is a complicated issue, because it obviously is essential, but, if review is not commissioned by the judiciary itself, it opens the door for outside interference with the judiciary and thus detracts from independence.

Turning to the country profiles, we focus on the highlights. It is not surprising that differences are large among countries given their recent histories. Tables 2 and 3 illustrate this for Denmark and Bulgaria.

---

\(^{6}\) Pearson correlation: 0.83 (N=24). This indicator of perceived independence by citizens is an average of three separate data sets. The correlation with these data sets separately is also high: 0.90 (N=26) with judicial independence in the Global Competitiveness Report of the World Economic Forum and 0.84 (N=22) with the impartiality of the criminal law system measured by the Rule of Law Index. Finally, the correlations with the European Barometer percentages of respondents that rate the independence of courts and judges as (fairly) good are 0.67 (N=24). Note that, while 23 judiciaries answered the questionnaire, 26 participated in the survey among judges (N is maximally 26).
Denmark shows a strong performance across the board with respect to objective and subjective independence and a strong, but less consistent, performance on accountability, while in Bulgaria objective independence and accountability are well arranged, but subjective independence scores are low. In Bulgaria many of the formal arrangements are state of the art, due to the need to rebuild judicial institutions after communism and due to the EC that demanded these arrangements as part of entry negotiations. There is a strong legal basis for an independent judiciary. Typical arrangements for this type of judiciary are random case allocation to reduce possibilities of corruption and the existence of mechanisms for external review of the judiciary. The perceptions of independence lag behind, which is likely to reflect reality, but to some degree may also have to do with slow adaptation of perceptions.

As to objective aspects, Denmark has an informal approach, showing in relatively low scores with regard to the constitutional position of the judiciary, case allocation, external activities of judges and reporting. This is more pronounced in the UK with its informal common law tradition, and even more so in Sweden and Finland. Table 4 gives the profile for Sweden. Weak formal arrangements of independence and accountability go together with positive perceptions in society and among judges about independence.
The profile of Bulgaria, which is similar to other countries in Eastern Europe, is not confined to Eastern Europe. Table 5 gives the profile of the Spanish judiciary, which scores low on subjective independence, although trust in the judiciary is higher than in the other branches of the state.

The results demonstrate that in order to improve outcomes more emphasis is required across Europe on subjective independence. The absence of data tells much about the importance attached to the opinions of court users; there should at least be systematic measurement of the views of court users to start to address this blind spot. From the views of the court users strategies can be developed to deliver justice in a way that citizens really want.

Figure 4. Indicators independence and accountability 2017, Sweden

Figure 5. Indicators independence and accountability 2017, Spain

The results demonstrate that in order to improve outcomes more emphasis is required across Europe on subjective independence. The absence of data tells much about the importance attached to the opinions of court users; there should at least be systematic measurement of the views of court users to start to address this blind spot. From the views of the court users strategies can be
developed to deliver justice in a way that citizens really want. From an analytical point of view it is striking that there seems to be no relation between objective and subjective independence. High levels of formal independence and accountability can go together with a high level of subjective independence (Denmark) and a low level (Bulgaria, Spain) and a low level of formal independence with a high level of subjective independence (Sweden) and a low level (Portugal). This finding was also noted in the literature discussed in Section II(2) and seems to be robust. One explanation was already hinted at: negative perceptions may trigger the response to strengthen formal safeguards. Another explanation has to do with the different recent histories of nations. Both explanations imply that strengthening formal independence can help to improve subjective independence. The outcomes are relevant at the national level and at the European level. In the next Section we will illustrate these uses.

VI. USE OF THE OUTCOME OF THE INDICATORS

1. Use at National Level: Example of the Netherlands

Since the results have only recently become available, the actions of Councils and other governing bodies, let alone the results of these actions, as a response to the report still have to follow. We take recourse to the first comprehensive measurement of the indicators in 2014/2015, albeit that the system of indicators was then still being developed. Figure 6 gives the country profile of the Netherlands. The profile has led the Council for the judiciary to a critical evaluation of all aspects covered by the indicators, and a selection of issues that had the highest priority and chance of success. It was concluded that in the area of independence the weak legal basis of the judiciary is an important issue, as the constitution does not guarantee the independence of the judge. The other issue is the appointment of members of the Council: the current mechanism does not meet the ENCJ guidelines of judicial members being elected by their peers (compliance with the ENCJ guidelines is part of indicator 2). As to accountability, the mechanism for case allocation received priority. All three weaknesses are currently addressed. In a parallel development the mechanism for case allocation received priority. All three weaknesses are currently addressed. In a parallel development the mechanism for case allocation received priority. All three weaknesses are currently addressed.

The profile has led the Council for the judiciary to a critical evaluation of all aspects covered by the indicators, and a selection of issues that had the highest priority and chance of success. It was concluded that in the area of independence the weak legal basis of the judiciary is an important issue, as the constitution does not guarantee the independence of the judge. The other issue is the appointment of members of the Council: the current mechanism does not meet the ENCJ guidelines of judicial members being elected by their peers (compliance with the ENCJ guidelines is part of indicator 2). As to accountability, the mechanism for case allocation received priority. All three weaknesses are currently addressed. In a parallel development the mechanism for case allocation received priority. All three weaknesses are currently addressed. In a parallel development the mechanism for case allocation received priority. All three weaknesses are currently addressed. In a parallel development the mechanism for case allocation received priority. All three weaknesses are currently addressed.

Figure 6. Indicators Independence and Accountability 2014/2015, Netherlands


70 This profile is not fully comparable with the recent profiles presented in Section IV due to improvements in the indicators and the methodology.

71 Proposal of law 34517.
2. Use of the Indicators at European Level: Risks to judicial independence

The outcomes of the indicators and the patterns they produce provide insight into the functioning of judiciaries in their democratic context. From the perspective of courts striving for independence, important risks for judicial independence are diminishing popular and governmental or political support. In diverse countries, governments and politicians criticise the judiciary either for electoral or policy reasons. There is, perhaps, decreasing self-restraint by both the media and politicians. Attacks have the effect, intended or otherwise, of discrediting the judiciary. In some cases, attacks are aimed at bringing the judiciary under the control of the other state powers. The indicators identify weaknesses in independence and accountability that make judiciaries vulnerable to such attempts. In the previous Section we identified two situations that pose risk.

- Strong formal arrangements that go together with negative perceptions in society of the judiciary and its independence. This situation makes the judiciary vulnerable to attempts by governments to remove formal protections and to intervene in the judiciary. Hungary and Poland are cases in point.
- Weak formal arrangements that go together with positive perceptions of the judiciary. If such judiciaries (mainly in Northern Europe) were to start losing popular support, the weak institutional arrangements offer limited protection against direct interference.

In both situations maintaining and strengthening popular support for the judiciary is important. While the determining factors of support are not well known, it is more than likely that excellent performance of the courts in the form of timely, impartial and well-reasoned decisions is helpful. This requires the courts to focus on their users, which they do not always currently do, as we found in Section V. Striving for accountability to society is a prerequisite. The findings of Section V can then give guidance on what to do. As to underlying motivations, the ENCJ has noted that an important risk to the independence of both the judiciary as a whole and the individual judge is the failure of judges to reflect changes in civil society, and their being out of touch with ordinary citizens.

Sufficient funding of the judiciary is also a pre-requisite to an excellent service to court users. In the previous Section we saw that the funding systems and the court management are the weakest aspects from the perspective of objective independence. Many judiciaries in Europe are entirely dependent on arbitrary decisions of government and are, as a result, underfunded. Furthermore, in allocating the budget, they are often dependent on court management by or under the control of the Minister of Justice. Finally, as to criticism of judicial decisions by the media, politicians, parliamentarians and the executive that is felt to be gratuitous by the judiciary, the courts can do more to strengthen their relationship with the press. Section V (Accountability, indicator 4) shows that, while some judiciaries are already effective in this area, there is generally room for improvement.

VII. DISCUSSION

The ENCJ developed a set of indicators to assess the state of independence and accountability of judiciaries, based on the notion that the two concepts were linked. Independence without accountability will not be accepted by society. Accountability without independence is bureaucratic decision-making. It needs to be emphasized that accountability to society is interpreted in a broad sense. It deals primarily with the transparency of procedures such as case allocation and withdrawal and recusal of judges and the transparency of judicial decisions. Only a small part of the accountability indicators is about managerial accountability regarding production and efficiency. In our view this interpretation reduces the tension between independence and accountability, as is regularly put forward in the literature. As to independence, the set consists of indicators of objective independence (formal, mostly legal, characteristics) and subjective independence (perceptions). The outcomes confirm that both aspects of independence are essential; objective and subjective independence often diverge in the profiles of individual countries. It cannot be taken for granted that countries that adopt best practices for formal judicial independence safeguards will achieve high levels of perceived independence. Conversely, it cannot be assumed that high levels of perceived independence will last forever so that strong formal arrangements are not necessary.

As to the way forward, making a systematic assessment of the level of independence and accountability achieved in practice by national legal systems is a crucial starting point for improving justice systems across the EU. Justice systems do not exist to benefit judges; they exist for the benefit of every citizen. Shining a light on systems that lack either or both of independence and accountability should enable their shortcomings to be more appropriately addressed. From a research perspective, this approach helps to make independence and accountability measurable, taking into account the complexity and many dimensions of both concepts. At the same time it must be recognized that the methodology contains (inter)subjective elements, for instance with regard to the scoring of the indicators, and it is based on self-evaluation which could make outcomes dependent on the incentives of those who conduct the evaluation. The latter aspect should not be a risk, as the indicators are about the formal arrangements in a country and otherwise observable phenomena, and can be readily checked by any knowledgeable observer. Still, there are no guarantees.

---

72 Poland is the most recent case in point. Hungary and Turkey went before. The criticism of the UK judges who decided the case on triggering the Brexit process is also relevant.
73 Poland, see EC, fn 14.
74 ENCI 2014, p40. See fn 44.
In the coming years, the ENCJ and its members and observers will, on the one hand, use the lessons that can be drawn from the country profiles to improve the functioning of the specific judicial systems and, on the other hand, seek external scrutiny of the methodology and outcomes of the indicators. It will also focus on getting information about the perceptions of court users on independence as well as accountability. Also, other ways to measure the ‘de facto’ independence of the courts will be considered. The latter aspect requires, in particular, further research.

To conclude, it is hoped that the move from debating the theoretical importance of judicial independence to the development of a practical analytical method for measuring it will enable justice systems across Europe better to uphold the rule of law and to improve the services they deliver to their citizens. At the same time, the method provides a more solid basis for research on independence and accountability.
For three decades, Tyler has been delivering solutions that transform the justice community. Our end-to-end Odyssey case management solution includes ODR, e-filing, and touch screens for judges — and is improving efficiency in courts, serving more than 120 million people around the world. Tyler empowers 15,000 government offices across Australia, Canada, the Caribbean, the U.K., the U.S., and other international locations.

Learn more at tylertech.com.
Viewing Judicial Independence and Accountability through the “Lens” of Performance Measurement and Management

Ingo Keilitz

Abstract
This article is a review of the European Network of Councils for the Judiciary’s (ENCJ) framework and vision of the independence, accountability, and transparency of the judiciaries in member states of the European Union. Its purpose is to aid ENCJ in the further development of its indicators of judicial independence and accountability. The focus of the article is on performance measurement and management (PMM) as seen, depending on one’s views, either as an instrument for strengthening judicial independence or, alternatively, an instrument for reining in the power of the judiciary and threatening its independence. The article begins with a general discussion of judicial independence, accountability, and transparency as seen through the “lens” of PMM. It continues with a critical review and assessment of the conceptual framework of the ENCJ’s 22 indicators and 64 sub-indicators of judicial independence and accountability and identifies several shortcomings that decrease the utility of the framework. It urges a rethinking of the conceptual framework and proposes an alternative model – an input/output/outcome “logic model” -- more amenable to understanding and improving indicators of judicial independence and accountability. It makes four recommendations aimed at a better alignment of ENCJ’s framework of indicators with principles and practices of modern PMM. The article concludes with a warning about troubling developments at a higher level of governance and politics that some see as a retreat from democracy and the rule of law in Europe and in many other parts of the world, one that poses an existential threat to the judiciary as a coequal partner in government.

1. Introduction
The goal of judicial independence is to secure for individual judges, courts, and court systems the independence to resolve disputes according to the law and to shield them from improper interference from the other branches of government, or private or partisan interests. A judge, court, or court system compromise their impartiality or independence, for example, when they merely ratify plea bargains, serve solely as a revenue-producing arm of government, or perfunctorily place their imprimatur on decisions made by others. Our understanding of judicial independence today encompasses not only control and authority over the legal decisions of individual judges but also an array of administrative powers of courts and judiciaries as organizations, including authority over budgeting, information technology, human resources, allocation of judicial services (supply chain management), judicial selection, retentions, and assignment, and the education and training of judges and justice system staff. As noted by van Dijk and Vos in their lead article in this special issue of the Journal, “insufficient and arbitrary funding of the judiciary, can make individual independence an empty shell.” Well designed and executed, judicial independence instills legitimacy and public trust and confidence in the judicial system of a country.

In 2013, the European Councils of the Judiciary (ENCJ), an organization of national councils of the judiciaries in the member states of the European Union (EU), began an ongoing initiative to evaluate the judicial independence and accountability of the judicial systems in the member states of the European Union (EU) through the “lens” of performance measurement. To date, ENCJ has

1 Ingo Keilitz, Ph.D., heads CourtMetrics, a consultancy in Williamsburg, Virginia, United States specializing in performance measurement and management of justice systems. He is a Research Associate of the Global Research Institute and a Visiting Scholar at the Public Policy Program at the College of William and Mary, Virginia, United States, and a Fulbright Scholar.

2 Our modern understanding of judicial independence and the separation of powers also includes judicial review, the power of the judiciary to correct the actions of the legislative and executive branches when they exceed their authority, such as occurred in Romania in July 2018 when the country’s constitutional court sided with the executive in the dismissal, ordered by the justice minister, of Laura Kovesi, the appointed head of the National Anticorruption Directorate (DNA) who had overseen the conviction of more than 1,000 business people and politicians, including nine former ministers. Legislation gave the DNA authority over high- and medium-level corruption cases. DNA is an independent entity in relation to the other judiciary branches, the prosecutor’s offices attached to these branches, as well as in relations to other public authorities. Corruption in eastern Europe: An enemy of crooked politicians is sacked for being too good at her job. The Economist, July 21, 2018, 38, https://www.economist.com/europe/2018/07/19/an-enemy-of-crooked-politicians-is-fired-in-romania [accessed July 30, 2018].


developed a framework and vision of independence and accountability and a set of indicators to evaluate the state of independence and accountability of EU judicial systems.

1.1. Purpose
This article is a review and an assessment of ENCJ’s framework and vision of the independence, accountability, and transparency of the judiciaries in member EU states. Its purpose is to aid ENCJ in the further development of its indicators of judicial independence and accountability. The focus of the article is on performance measurement and management (PMM) as seen, depending on one’s views, either as an instrument for strengthening judicial independence or, alternatively, an instrument for reigning in the power of the judiciary and threatening its independence.

The article explores how ENCJ “frames” judicial independence by its identification, classification, definition, and use of the 22 indicators and 64 sub-indicators of judicial independence and accountability as described in its 2017 report. Because the focus of the article is on the conceptual framework of the indicators, which is applied consistently across the all the indicators and sub-indicators developed by ENCJ, and not on the detailed substance of the indicators, the exploration is limited to the 22 indicators and 64 indicators of independence and accountability, and does not cover the 21 new indicators and 82 sub-indicators of the quality of justice that are still in the early development stages in ENJC’s ongoing initiative. For the same reason, the article does not delve into the substance of the sub-indicators.

Key questions and related issues addressed by this article are posed in italics in this and the next paragraph. Do the demands of rigorous PMM as part of the administration of justice – including the regular and continuous monitoring of performance in a transparent and accountable manner – threaten judicial independence? It is inevitable that there will always be defenders and critics of the performance of courts (e.g., about delays, costs, and reliability) and arguments over how to improve court organization and operations, but how can PMM foster in the judiciary the strength to resist encroachments by the other branches of government, private or partisan interests into its domain?

What is the appropriate scope of judicial independence at the level of the individual judge (decisional independence) and the organizational level of a court or a national court system (institutional independence)? Where should the lines be drawn that separate the powers of the judicial, executive, and legislative powers of government and that protect judicial independence? Does judicial independence mean complete autonomy and self-government, and if not, how much comity, coordination, and cooperation with the other branches of Government and other public and private interests does it require? Most observers would agree that judicial independence does not demand absolute autonomy or isolation of judges and courts. Do the demands for measurement and management of the performance from within (e.g., emanating from presiding judges and court administrators) and outside the judiciary constitute an imposition that threaten judicial independence? These are fraught questions faced by judiciaries in Europe and around the globe.

Much like the relationship between doctors and hospital administrators, the relationship between judges and court administrators (including judges in administrative positions) -- who generally are more likely to advocate for rigorous performance measurement -- are strained and fraught, more so than they seem to be in Europe than in the United States. Both doctors and judges are semi-autonomous professionals who often chafe at attempts by administrators to limit their freedom to do as they please (e.g., doctors order medical tests that are not needed and judges postpone trials too many times contributing to court delay).

In the United States, a recognized world leader judicial administration, the field of court administration is more mature and more stable than in Europe. All the 50 states in the United States have state court administrators who have varying degrees of management control over the courts in their states. Almost all courts at all levels employ court administrators or court executives, many with considerable influence over judicial administration. At the national level, professional organizations such as the Conference of Chief Judges, the Conference of Court Administrators, the National Center for State Courts, and the National Association for Court Management promote the role of court administration in governance of the judicial branch.

While it is common for judges and judicial groups in the United States to complain about an infringement of the independence of the judiciary and a blurring of the separation of powers, such as when court administration pressures courts to adopt rigorous, transparent, and accountable performance measurement and management, or to cut costs, the complaints have been more muted and have not created the level of tension and rift between judges and court administration that we see in Europe.

1.2. Contents
Following this Introduction, this article is organized into four sections. Section 2 is a discussion of judicial independence and judicial accountability as seen through the “lens” of performance measurement and management (PMM). Section 3 is a critical

---

5 ENCJ does not specifically define “indicators.” The different meaning and definition of the terms “indicator” and “measure” are perennial topics of much discussion. They terms are used here interchangeably largely as matter of convenience because ENCJ uses the term “indicator” and I have preferred the term “measures” in my research and writing.


review and assessment of the conceptual framework of the ENCJ’s 22 indicators and 64 sub-indicators of judicial independence and accountabilit

2. Performance measurement, accountability, and transparency

Beginning with a definition of PMM, this section is a discussion of judicial independence and judicial accountability as seen through the “lens” of PMM.

2.1. Definition of performance measurement and management

There is no precise agreed-upon meaning of “performance measurement.” For this article, performance measurement and management (PMM) in courts and justice systems is defined as it is in the second edition of the International Consortium of Court Excellence’s International Framework of Court Excellence:

“Court performance measurement and management (PMM) is the discipline and the process of monitoring, analyzing, and using organizational performance data on a regular (ideally in real or near-real time) and continuous basis for the purposes of improvements in organizational efficiency and effectiveness, in transparency and accountability, and in public trust and confidence in the courts and the justice system.”

Several parts of this definition merit highlighting. First, performance measurement is defined as a management discipline encompassing a system of concepts, methods, and techniques, as well as a process. The definition of PMM as a discipline addressing the key question “How are we performing?” emphasizes that PMM is a rigorous way of thinking about solving problems, discovering opportunities, and identifying possible solutions (whatever moves the measurement “needle”). Second, by recognizing that performance measurement data are of no use if not used, measurement is explicitly paired with performance management to emphasize that to be used effectively PMM should be infused into the very DNA of governance and management of operations such as budgeting, resource management, supply chain management, and strategic planning. This pairing, which firmly anchors the discipline of PMM in the tradition of management and business, is a relatively new development widely seen as a major step in transforming measurement into management for real organizational change.

Third, the discipline of PMM should be practiced on a regular and continuing basis, ideally in real time or near real-time, as performance occurs and as the needs of users of specific measures dictate. Actionable performance measures are sensitive to interventions and, therefore, must be taken and used in tight timeframes, not months or years after measurement occurred. PMM is analogous to the “measurements” we take as we monitor the performance of our car using the car’s dashboard of indicators. It would be nonsensical and unsafe if our car’s speedometer only registered speeds once every hour. Finally, the definition aims PMM...
toward specific purposes including efficiency and effectiveness, transparency and accountability, and public trust and confidence in the judicial branch.

2.2. Truth, accountability, and transparency

In his inspiring 2017 book, *Principles*, Ray Dalio urges us to trust in "radical truth and radical transparency" in our work in organizations.14 "Understanding what is true is essential for success, and being radically transparent about mistakes and weaknesses, helps to create the understanding that leads to improvements," he writes.15

To its credit, the ENJC has approached its work on judicial independence, accountability, and transparency through the "lens" of PMM. ENCJ has an understanding that superior performance is the product of accountability. It recognizes that judicial independence and performance accountability and transparency go hand in hand, the latter being a necessary condition of the former. In its 2017 report, ENCJ asserts that a "[j]udiciary that does not want to be accountable to society and has no eye for societal needs will not gain the trust of society and will endanger its independence in the short or long run.") Conversely, it notes that "accountability without independence" reduces a judiciary to an agency of the executive or legislative branches.16 This position is reinforced by van Dijk and Vos, noting that justice systems "that are accountable to no one are likely to have weaker incentives to improve themselves than systems that open up to the outside world and tackle criticism seriously. By emphasizing and strengthening independence solely judiciaries risk insulating themselves from society and becoming irresponsible to justified demands of society."17

Because of their "independent" status, the judiciaries' rules and procedures are not based on the same analysis and subject to the same level of scrutiny as those of agencies of the executive and legislative branches. Many, if not most, of the judges in the Netherlands and elsewhere in Europe and beyond generally embrace -- or at least acquiesce to -- the need for accountability and transparency in theory. They recognize that judicial reform aimed at making courts and other judicia institutions more efficient and effective nearly always involves trade-offs between independence and accountability. Perhaps, as a nod toward comity, coordination, and cooperation with other branches of government, and other public and private interests, they acknowledge that the greater degree of freedom from scrutiny and interference traditionally enjoyed by courts compared to other public institutions (at least in much of the Western world) must be balanced by the *quid pro quo* of transparency and accountability. Some see their accountability and transparency as a source of pride reflecting their commitment to serve the people which in turn, they believe, ensures the public trust and confidence.

However, to many judges the words "accountability" and "transparency" conjure up something to be feared. They have grave concerns about judiciaries' acceptance of rigorous measurement and management of performance enabling such accountability and transparency.18 They feel pressured by executive branch representatives (including ministries of justice) and court management to balance independence with accountability implemented by performance measurement focused on, for example: (a) increasing the productivity of judges and courts (i.e., resolving more cases more quickly); (b) reducing costs by closing or consolidating courts; (c) rationalizing and justifying the allocation and distribution of judicial resources including judges, administrative staff, and courts; and, generally, (d) improving supply chain management (i.e., balancing the supply of judicial services delivered to people with the demand for those services). Courts may be ill-equipped with limited staff and technical resources to translate concepts and roomy objectives and definitions of independence, accountability, and transparency into precisely defined performance measures and metrics. Judges perceive these as constraints or threat to their independence.

On the other hand, advocates of performance accountability and transparency -- both within and outside of judicial institutions -- argue that, done well, performance measurement shines a light on real progress in solving entrenched problems that matter to ordinary citizens. The promise of organizational performance measurement is suggested by maxims that have gained much currency in recent years: *You can’t manage what you can’t measure. What gets measured gets attention. What gets measured gets done. Measure what matters, count what counts. Measure what you treasure.*19

In an essay in the *Wall Street Journal* in 2013, accompanied by a large picture of a globe wrapped in a tape measure, Bill Gates, the former head of Microsoft and co-chair of the Bill and Melinda Gates Foundation, wrote that his plan to solve the world’s biggest problems was, plain and simple, to measure them. "In the past year," Gates wrote, "I have been struck by how important measurement is to improving the human condition. You can achieve incredible progress if you set a clear goal and find a measure that will drive progress toward that goal. This may seem basic, but it is amazing how often it is not done and how hard it is to get

---

15 Ibid., p. 323.
16 ENCJ Report, 2016-2017, supra note 6, p. 11.
17 van Dijk & Vos, supra note 3, pp. 11-12.
right.” In a similar vein, John Doerr, in his 2018 book, *Measure What Matters: How Google, Bono, and the Gates Foundation Rock the World with OKRs* (OKRs are objectives and key results) echoes Gates’ view. He exhorts private and public organizations to embrace performance measurement and management as a “sharp-edged tool for world-class execution.”

3. The conceptual framework of ENCJ’s indicators of judicial independence and accountability

This section presents a critical review of the conceptual framework of ENCJ’s indicators of judicial independence and accountability. As noted in the Introduction, the focus is mainly on the conceptual framework and not the substance of the 22 indicators and 64 sub-indicators of independence and accountability. Following the critical review, which finds the conceptual framework problematic, the section offers an alternative model – a logic model -- that is simpler, aligns with the principles and practices of contemporary PMM, and promises a more productive platform to launch further development of ENCJ’s indicators of judicial independence and accountability.

3.1 Classification and categorization of the indicators

ENCJ classifies a total of 22 indicators and 64 sub-indicators -- 13 indicators and 33 sub-indicators of judicial independence and nine indicators and 31 sub-indicators of judicial accountability -- into four categories: (a) “objective” or “formal” independence and accountability; (b) “subjective” or perceived formal independence and accountability; (c) individual judge’s judicial independence; and, (d) the independence of the judicial system as a whole. This classification and categorization can be depicted as 2 X 2 matrixes as shown Table 1 and Table 2 below. (The 22 indicators in the tables are identified only by name and only by the number of sub-indicators associated with each of the indicators that appears in parenthesis after each indicator.)

The indicators are described as “key aspects” of independence and accountability “generally consisting of sub indicators to capture the diverse elements that are relevant for a key aspect.” For example, Indicator No. 1, in the upper left quadrant of Table 1, is identified with the key aspect of the “legal basis of independence,” consisting of five sub-indicators including: (i) formal guarantees of the independence of the Judiciary; (ii) formal assurances that judges are bound only by the law; (iii) formal methods for the determination of judges’ salaries; (iv) formal mechanisms for the adjustment of judges’ salaries; and, (v) formal guarantees for involvement of judges in the development of legal and judicial reform.

Table 1. Thirteen ENCJ Indicators and 33 Sub-Indicators of Judicial Independence in Three Categories

<table>
<thead>
<tr>
<th>Objective/Formal</th>
<th>Subjective/Perceived</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Legal basis (5 sub-indicators)</td>
<td>9. Public perception (3)</td>
</tr>
<tr>
<td>2. Organizational autonomy (4)</td>
<td>10. Trust in Judiciary (1)</td>
</tr>
<tr>
<td>3. Funding (4)</td>
<td>12. Public perception of corruption (1)</td>
</tr>
<tr>
<td>4. Management (1)</td>
<td>12. Court users’ perceptions (1)</td>
</tr>
<tr>
<td>5. Human resource decisions (5)</td>
<td>13. Judges’ perception (1)</td>
</tr>
<tr>
<td>6. Disciplinary measures (2)</td>
<td></td>
</tr>
<tr>
<td>7. Non-transferability (2)</td>
<td></td>
</tr>
<tr>
<td>8. Internal independence (3)</td>
<td></td>
</tr>
</tbody>
</table>


22 van Dijk & Vos, supra note 2, p. 9-12 devote a large portion of their article to the identification and explanation of the sub-indicators. With the exception of noting the numbers of sub-indicators associated with the classification and categorization of the 22 indicators in Section 3, this article does not delve into the substance of the sub-indicators.
Table 2. Nine ENCJ Indicators and 31 Sub-indicators of Accountability in Two Categories

Nine ENCJ Indicators and 31 Sub-indicators of Accountability

<table>
<thead>
<tr>
<th>Objective/Formal</th>
<th>Subjective/Perceived</th>
</tr>
</thead>
<tbody>
<tr>
<td>Judiciary</td>
<td></td>
</tr>
<tr>
<td>1. Case allocation (2 sub-indicators)</td>
<td>None*</td>
</tr>
<tr>
<td>2. Complaints procedure (5)</td>
<td></td>
</tr>
<tr>
<td>3. Reporting (4)</td>
<td></td>
</tr>
<tr>
<td>4. Relations with Press (3)</td>
<td></td>
</tr>
<tr>
<td>5. External review (2)</td>
<td></td>
</tr>
<tr>
<td>Judge</td>
<td></td>
</tr>
<tr>
<td>6. Code of ethics (3)</td>
<td>*But indicators and sub-indicators of independence are relevant in this category.</td>
</tr>
<tr>
<td>7. Withdrawal and recusal (5)</td>
<td></td>
</tr>
<tr>
<td>8. External Activities (4)</td>
<td></td>
</tr>
<tr>
<td>9. Understandable procedures (3)</td>
<td></td>
</tr>
</tbody>
</table>

Indicators of objective or formal independence and accountability, in the upper-left quadrants (for the judiciary as a whole) and in the lower-left quadrants (for individual judges) in Table 1 and Table 2, are conceived as “formal, legal safeguards,” i.e., “actual” resources, processes, and formal arrangements that can be directly observed and, presumably, counted and measured.

In public sector administration, objective data or, more accurately, “quantitative data,” is viewed as the gold standard because it is based on measurable facts that are presumed to reflect the “real” world that can be observed, counted, and measured. As is explained in Section 4, the ENCJ indicators and sub-indicators do not meet this standard. An example consistent with this gold standard of quantitative data might include measures such as the number of complaints about judges registered, or the proportion of judges disciplined or dismissed. Objective and quantitative measures such as these can be aggregated to reflect the judiciary as a whole (e.g., the proportion of all dismissed or involuntarily relocated judges in a country, region, or court dismissed) and disaggregated to show the performance of individual judges (e.g., the number of complaints made against them).

Subjective independence or perceived independence in ENCJ’s scheme (the term “subjective” is used interchangeably with the term “perceived”), depicted in the left halves of Table 1 and Table 2, is conceived as perceptions of independence and accountability in society, including those of the general public, users of courts, private companies, experts of various stripes, and judges themselves. The data is collected through participant observation and understood from the respondent’s perspective. For reasons not altogether clear, the five indicators of judicial independence in the left column of Table 1 make no distinction between indicators for the judiciary as a whole and individual judges. There are no subjective/perceived indicators for judicial accountability in Table 2.

Subjective data are based on personal opinions, assumptions, interpretations, and beliefs gathered by surveys or reports of groups of experts. For example, the Eurobarometer Flash survey of judicial independence conducted by the European Commission, between 25 and 26 January 2017, the results of which are part of Indicator No. 9 in ENCJ’s scheme, relies on survey methodology including questions such as: “Could you tell me to what extent each of the following reasons [e.g., no interference or pressure from government and politicians] explains your rating of the independence of the justice system [in your country]? Respondents are asked to rate their agreement along a five-point scale including “very much,” “somewhat,” “not really,” “not at all,” and “don’t know.”

### 3.2. Shortcomings of the ENJC’s scheme

The ENCJ’s classification and categorization scheme has shortcomings that diminish its understanding and usefulness.

#### 3.2.1 Objective and factual versus subjective and perceived

First, the distinction between objective and factual, on the one hand, and subjective and perceived, on the other, is blurry. Most of the indicators and sub-indicators are based on perceptions as measured by survey responses. However, as categorized in ENCJ’s scheme, subjective independence is measured by only five indicators (Indicators Nos. 9 through 13) of perceived independence in response to surveys of both the judiciary as a whole and individual judges. No distinction is made between subjective/perceived indicators of independence (see Table 1 above) and no subjective perceived indicators are identified for judicial accountability (Table

---

2). The difficulty is that the indicators of objective/formal independence also are generated from the opinions and perception of surveyed experts. 24 A similar situation holds for judicial accountability. Moreover, the ENCJ describes the indicators in the European Union’s Justice Scoreboard and the other third-party surveys it relies on as “objective,” 25 when in fact those indicators also are based on independence as perceived, i.e., “subjective,” using ENCP’s definition of this term. Putting the indicators in the objective category does not make them so.

In their lead article in this special issue of the Journal, van Dijk and Vos note that the “scores for the indicators of objective independence and accountability were derived from the questionnaire, whilst those for subjective independence were based on international opinion surveys,” suggesting that the former are somehow based on objective data. 26 In fact, the former also are derived from surveys based on the perceptions and opinions of judges, members of the councils for the judiciary, and other individuals in the governing bodies such as ministries of justice. Van Dijk and Vos do recognize that this may lead to self-serving bias. While many of the indicators are observable things (e.g., the existence of laws, regulations, and rules for the adjustment of judges’ salaries), and can be readily verified by any knowledgeable observer, most are subject to considerable variations of opinions and possible bias (e.g., the sufficiency of budgets). As further explained below, the circumscribed use of survey data based on perception is valuable; a complete reliance on opinions is not.

A long-standing problem with survey and polling data is that in professional circles the term “subjective” is often couched in pejorative terms meant to characterize subjective measures as “soft,” mere opinion, not empirical, and not based on the application of the scientific method. This is unfortunate if by “subjective” we mean simply that the data is collected through participant observation and understood from the respondents’ perspectives. 27 Court user satisfaction measures such as the percent of court users who believe that the court they have just exited provides procedural justice (i.e., accessible, fair, accurate, timely, knowledgeable, and courteous judicial services), are among the most robust and widely used measures by justice systems around the globe. 28 Because perception data reflect personal ideology, intuition, and even personal bias, experts in public sector performance generally prefer objective or quantitative data. However, this preference is deterministic only if no good objective data are available or (and this is an important distinction) when the beliefs, opinions, and feelings, including the preferences and biases of respondents, are precisely what is sought, as is the case in measuring “procedural justice” or “procedural fairness” as experienced by users of courts. 29

3.3.2 Organizational versus individual judges’ performance

Another shortcoming might be better characterized as a problem with the placement of two separate processes and disciplines into the same conceptual framework – organizational performance measurement of the judiciary as whole and judicial performance evaluation (JPE) of individual judges. The U.S.-based National Center for State Courts has developed extensive separate resource guides for organizational (court) performance measurement 30 and JPE 31 that reflect the differences between these two disciplines and processes. Business leaders have learned, painfully so according to John Doerr, that organizations and individuals should be measured and managed quite differently. 32

Effective court performance measurement, anchored by the right performance measures, indicates organizational performance, and not necessarily the performance of individual judges or court staff. There is evidence to suggest that highly competent judges and court staff do not necessarily guarantee a high-performing court. And a court with the best practices supported by good supporting infrastructure and technology can operate at high levels with less than highly competent judges and staff, though not for the long-term. Organizational performance measurement and individual performance evaluation differ in purpose, methodology, interpretation, and use. The evaluation of individual judges with personal problems such as disabilities, health issues, personality disorders, marital problems, alcoholism, or drug addiction simply have no parallel in organizational performance management.

Reporting the performance of individual judges as part of organizational court-wide or justice system-wide performance is likely to cause animosity and divisiveness. It can carry the danger of derailing an otherwise promising performance measurement effort. Disaggregations or breakdowns of organizational court performance to the individual judge level figure prominently in the reasons raised in opposition to organizational performance measurement. 33 There is today simply insufficient knowledge and experience to support a formal linking of the two. For example, a court that is “speedy” – a court timely and expeditious in its case processing

---

24 See Van Dijk and Vos, this issue, need page citation
26 Van Dijk and Vos, supra note 2, page citation.
27 There is a fundamental and clearer distinction to be made between two types of data: qualitative and quantitative. Data is “quantitative” if it is in numerical form and “qualitative” if it is not.
28 The measure is part of the most prominent models of performance measurement and management (PMM) in the world including the Global Measures of Court Performance, the CourTools, and the European Commission for the Efficiency of Justice (CEPEJ).
29 Cf. Measure 1, Court User Satisfaction, Global Measures, supra note 13, p. 25.
32 See Chapter 15, Doerr, supra note 21.
33 Keilitz, supra note 18.
is generally considered a “good” court; a court with a large backlog of cases awaiting resolution wherein people experience considerable court delay is generally viewed as a court in need of improvement. No such consensus exists about a “speedy” judge. Many so called “slow” judges, are considered deliberative, thoughtful, and thorough, and are held in high esteem by their colleagues for their integrity, opinion writing, mentoring of junior judges, and management and leadership skills.

Consistent with the literature on JPE, it might be most prudent that organizational performance data be limited to providing only a background and broad context for the discussion of individual performances of judges and staff, not to evaluate, to set individual goals, and not to rate and rank judges against their peers on measures such as clearance rate. For example, a judge could be made aware that he or she is resolving cases at a clearance rate lower than the court and lower than the overall average of the courts in a jurisdiction.

3.2.3 Cohesion and Logical Connection

A final shortcoming of ENCJ’s conceptual framework of indicators and sub-indicators, which is related to the above, is that the classification and categorizations of indicators and sub-categories lack coherence and logical connections to principles and concepts of performance measurement. Why, for example, is there a distinction between objective and subjective data even made? What purpose does it serve? Is it helpful in identifying further development of the indicators and sub-indicators? Users of the framework may be prone to be confused by its complexity, especially in the context of the sheer number of indicators – a total of 43 indicators and 146 sub-indicators when one includes the 21 new indicators and 82 sub-indicators of the quality of justice that are still in the development stages. The issues that this shortcoming raises are touched upon in the following subsection and directly in Section 4, “Principles and practices of modern performance measurement and management.”

3.3. A logic model of measures of judicial independence and accountability

The figure below depicts a logic model (also called an outcome-sequence chart) adapted here to identify and to prioritize indicators of judicial independence and accountability. It is offered here as an alternative to ENCJ’s conceptual framework, one that promises to make up the shortcomings of the framework identified in the previous section. The model is much simpler, easier to understand, and to apply. In common sense terms, the elements or aspects of independence of accountability are categorized simply as the things we have at our disposal (inputs), the things we do (outputs), and the things we accomplish (outcomes).

The logic model, especially when depicted as in the accompanying figure, provides a visualization of what an indicator is intended to measure. It prompts fundamental questions such as: What is it that is being measured? Why does it matter and to whom? What critical outcomes and achievements do the things we have (inputs) and the things we do (outputs) lead to? Using the inputs of formal safeguards identified in the EMCJ indicators – i.e., one of the set of the things we have in the logic model – to highlight the fundamental questions that the logic model prompts us ask and think about, consider what we might count if we want to measure the success of doctors, hospitals, and the health care system, or the success of teachers, schools, and educational systems of the countries in the EU. What should we count? What should we measure that truly matters? What is the logic to support it? Few health care and education leaders and managers in these sectors, as well as patients and students, would be satisfied, for example, with an outcome indicator such as formal assurances that doctors and teachers are bound by law modeled on one of the sub-indicators of Indicator No. 1, Legal Basis of Independence, in the objective independence of the judiciary category in the ENCJ framework.

Finally, the logic model has coherence and logical connections to principles, concepts, and sound practices of PMM. The descriptions of PMM in international models such as the Global Measures of Court Performance and the CourTools include a concise operational definition, statements of purpose, and rationales for the prescribed performance measures including their alignment with values and critical success factors. This supporting logic is a fundamental requirement of the scientific method. No logic, no science. I strongly believe that such logic needs to be included in the ENJC framework.

---

35 See note on Indicator 1 and associated sub-indicators in the text following supra note 22.
36 Global Measures, supra note 13.
3.4 Externally oriented versus internally oriented measures

Judicial independence and accountability are not ends in-and-of themselves, but rather means to ends that matter to the public. Justice systems do not exist for the benefit of judges but rather for the benefit of the citizens they serve. Indicators of internal inputs (such as available human and fiscal resources) and outputs (such as activities and procedures for disciplining judges) focused on justice system "insiders" running the courts may prompt outside observers to see the judiciary as a “closed shop” operating solely for its own benefit. On the other hand, outcome measures show how well or how poorly a court system is making things better for people by improvements in their status or condition, instead of how much effort (outputs) has been expended or by with what resources (inputs).

The European Commission and the Council of Europe, as well as other groups such as the International Consortium for Court Excellence and the National Center for State Courts in the United States, make distinctions between two types of justice system performance measures. The measures of the first type are seen as core judicial service performance measures, defined as externally oriented outcome measures that register changes in the status or conditions of the people served by the courts. In justice systems that have PMM systems, the former type tends to be limited to a vital few aligned with the handful of mission-critical success factors of courts such as access to justice, expedition and timeliness, fairness, and public trust and confidence in the courts.

Performance measures of the second type are internally oriented input and output measures of court administration, structures, organization, and operations. These measures can be quite valid and useful organizational, operational, structural (internal) measures – judiciaries need to know how many courts and judges are in their jurisdictions -- but they have limited value as core judicial service performance measures. Furthermore, we simply do not know if the inputs and outputs measured are directly correlated with outcomes. In fact, there is research by one of the authors of an article in this special issue of the Journal, showing that they may be negatively correlated. These internally oriented measures also can expand without limits to cover a myriad of internal input and output measures of court administration, structures, and operations that may be of special interest limited to only a few agents of the justice system or a special issue at a particular time.

All but five of the 22 ENCJ’s indicators are input or output measures of formal arrangements, not outcome measures, and all five are measures of perceptions. Again, simply put, inputs are the resources we have; outputs are the things we do; and outcomes are what we accomplish. This input-output-outcome scheme, highlights problems in the ENCJ framework of indicators recognized by van Dijk and Vos, i.e., the gap between de jure and de facto independence and accountability, the difference between what we would like to see happen and what is actually happening. They write: “This allows for the possibility that a legal system may possess all the formal requirements of independence, but nonetheless have judges who do not act independently. It allows also for the possibility that systems are not transparent, such that the public cannot determine whether the judiciary is independent or not.”

38 See, for example, European Commission for the Efficiency of Justice, (CEPEJ), Measuring the Quality of Justice: https://rm.coe.int/16807477e4 accessed August 1, 2018.

39 S. Voigt, S. & N. El-Bialy (2014), Identifying the Determinants of Judicial Performance: Taxpayers’ Money Well Spent? European Journal of Law and Economics, 11 December 2014, p, 41, finding that rates of case resolution are not a function of court budgets (i.e., higher budgets will not necessarily “buy” increased efficiency) and that resolution rates are negatively correlated with the existence of judicial councils https://link.springer.com/article/10.1007/s10657-014-9474-8 [accessed 2 August 2018].

40 Van Dijk and Vos, supra note 3, page citation needed
As it continues its development of indicators of judicial independence and accountability, I urge ENCJ to adopt and use the logic model along with the principles of modern PMM described in the next section.

4. Principles and practices of modern performance measurement and management

This section consists of four recommendations and supporting commentary aimed at a better alignment of ENCJ’s framework of indicators with principles and practices of modern performance measurement and management (PMM) including: (i) the emphasis on outcomes and measures of judicial independence and accountability that matter to the people that are served by the judicial system instead of the “insiders” who run the system; (ii) the prioritization and streamlining of the indicators and sub-indicators to a vital (and manageable) few; (iii) definitions of the indicators and sub-indicators that meet five “SMART” criteria – specific, measurable, achievable, relevant, and time-based; and, (iv) the management, not just the measurement, of judicial independence and accountability.

4.1. A preference for outcome measures

In their continuing efforts to improve judicial independence and accountability through the “lens” of performance measurement, ENCJ and its member judicial councils and their partners, should count what counts, measure what matters, and emphasize outcome indicators that matter to those who are served by the judiciary, instead of those who “run” the judiciary. They should maintain an external orientation and look outward instead of inward.

This recommendation has been alluded to in several places earlier in this article. It’s importance merits making it explicit here. Figuratively and literally, performance does not count unless it is related to the things that really matter, that are critical to the success – the efficiency, effectiveness, and accountability of a justice system. The connection between well-known health indicators like blood pressure, cholesterol level, and blood glucose, and the outcomes for our health and well-being is self-evident to most people. We know that these measures mean something vital and something very important to us. Such key success factors for judiciaries as access to justice, fairness, and timeliness have been referred to in the literature of justice system performance measurement as major performance areas, high level goals and objectives, standards of success, perspectives, domains, performance criteria, key results factors, and key outcomes. Whatever they are called, they form the framework of a judiciary’s accountability and transparency to the public and other stakeholders.

Increasingly, court systems’ stakeholders – ordinary citizens, politicians, litigants, legislators, jurors, witnesses, and court employees – are calling for clear evidence that the resources court systems expend bear a causal relationship to, or are at least are clearly associated with, benefits for people. Core outcome measures emphasize the condition or status of the recipients of judicial services or the participants in court programs (outcomes) rather than the resources (inputs) or the activities and processes, programs and activities (outputs). That is, they measure results and accomplishments, not merely resources, level of effort, and work performed. They measure ends rather than the means to achieve them. It is by this outcome orientation, for example, that the Global Measures help link the International Consortium of Court Excellence’s values and areas of court excellence referred to “drivers” (e.g., court management and leadership) and “systems and enablers” (e.g., court policies; human, material, and financial resources; and court proceedings) to results and outcomes.

4.2. A vital few indicators

ENCJ’s indicators and sub-indicators should be streamlined to a vital few instead of a “trivial” many, a vital few that constitute a “balanced scorecard” of measures of judicial independence and accountability.

The ENCJ should streamline the 22 indicators and 64 sub-indicators to a vital few core measures of judicial independence and accountability. This recommendation becomes more cogent in view of ENCJ’s plans to add 20 indicators and 82 sub-indicators of quality of justice – an overwhelming total of 82 indicators and 126 sub-indicators. (Moses made do with just ten commandments. Heaven help the poor Christians if they had to contend with 82 commandments and 126 sub-commandments.) Because most people can remember no more than three to seven things at any one time, too many measures scatter their attention and make them less efficient and effective. The key to a successful PMM is paring down the number of measures to a vital few that are clearly linked to critical success factors.

The question of the ideal number of performance measures that should be part of a dashboard of a car (speedometer, odometer, tachometer, and other instrumentation for monitoring temperature, gas, and oil pressure), for example, is informed by the idea of “a vital few instead of a trivial many.” There is only so much that a driver of a car can pay attention to while driving. The indicators should be the critical few gauges, instrument panel focused on a safe, efficient, and timely journey without unnecessary distraction.

As it moves forward with its important work, I urge the ENCJ to look at other prominent models of performance measurement and management, in particular: (a) the International Consortium for Court Excellence’s Global Measures of Court Performance, which

---

41 Global Measures, supra note 13.
42 The phrase “vital few versus trivial many” used in relation to the number of desirable performance measures is attributable to Mark Graham Brown (1996) in Keeping Score: Using the Right Metrics to Drive World-Class Performance. New York: Quality Resources.
43 Global Measures, supra note 13.
is part of the International Framework of Court Excellence of the International Consortium of Court Excellence; (b) the CourTools and Appellate CourTools44 of the National Center for State Courts; and (c) the 12 sub-goals and 21 indicators of Goal 16, the justice goal,45 of the United Nations Sustainable Development Goals (SDGs).46 These three models provide many insights into methodology of measuring independence and accountability including the need to limit the number of measures. In a 2016 paper,47 I joined a chorus of critics pointing out that Goal 16 and its associated 12 performance targets and 21 proposed provisional indicators of the United Nations SDGs, as well as the total SDGs package adopted by the United Nations in September 2015, with 17 goals and 160 sub-goals, was simply too much to handle. The Global Measures and the CourTools both have a streamlined number of measures, emphasize outcome measures, and are clearly linked to fundamental values and principles of justice administration.48

An effective PMM system should include only a handful of core measures49 that gauge important outcomes focused on key performance areas. It is this linkage with key performance areas that also helps to limit the number of core measures to a vital few insofar as broad strategic goals tied to citizens’ needs and wants tend to be relatively few in number. Another way of focusing measurement on a vital few performance measures is to build hierarchies (some call these “families” or “cascades”) of measures with the few core measures sitting at the top of a hierarchy of related subordinate measures. In this way users can “drill down” into ever more complexity only if they choose to do so. Creating measurement hierarchies keep things simple, allowing many measures in the system yet maintaining the focus on the vital few measures at the top of the hierarchy. Still another way is to develop a simple index of several measures, variables or aspects of a single area of performance, for example, expedition and timeliness of court case processing.50

4.3. Make ENCJ’s indicators SMART

The ENCJ should operationally define the indicators of judicial independence, which in their present form are no more than statements of goals or issues, and make the definitions concise, measurable, and actionable; i.e., make them SMART – specific, measurable, attainable, relevant, and time-bound.

Conceptual clarity and focus are requirements of good definitions of goals that make measurement of success possible. In this section, I argue that many, if not most, of the ENCJ indicators and sub-indicators are not measurable and manageable as currently formulated.

The concept of measurement means various things to different people across professions, fields, and disciplines. Generally, however, the scientific community sees measurement as a quantitatively expressed reduction of uncertainty based on one or more observations. Well formulated measures and indicators are SMART. That is, they are specific, measurable, attainable (actionable), relevant, and time-based. Conceived in this way, precise definition and effective measurement are necessary processes to reduce uncertainty.

In its papers, ENJC reflects some apprehensions about the challenge of measuring judicial independence and accountability. Is it even possible to measure something as complex and multifaceted as judicial independence? As pointed out by Douglas W. Hubbard in his book, How to Measure Anything,51 if we incorrectly believe that measurement means that we must meet some unachievable standard of certainty, then I would agree that judicial independence is immeasurable. But when scientists perform measurements they seem to be using a different definition in terms of observations that quantitively reduce uncertainty. To them measurement is a quantitatively expressed reduction of uncertainty based on one or more observations. Anything including judicial independence and accountability can be measured. If it matters at all, it is detectable and measurable, and it reduces uncertainty.

That said, the 22 indicators and 64 sub-indicators need to be developed into concise, measurable, and actionable performance measures. The ENCJ surely recognizes this. The popular mnemonic acronym “SMART” has been used to guide the setting of goals and objectives, and the formulation of performance targets, in the fields of management for more than 50 years.52 The following are the most common requirements of the SMART criteria:

Specific: focused, clear and unambiguous, targeting a specific area for improvement in a manner that is easy to understand.

Measurable: quantifies progress and success (e.g., in terms of how much and how many).

Achievable: realistic and actionable.

44 CourTools, supra note 37.
45 Goal 16: peace, non-violence, safety, and security; access to justice, and just and inclusive societies; and effective, inclusive, and accountable institutions.
48 It is uncanny that the predecessor measurement system of the CourTools (which contains only ten measures), that was part of the seminal 1997 Trial Court Performance Standards included 21 standards and 68 performance measures, almost the same number of the ENCJ framework.
49 A core performance measure is a primary indicator – like the speedometer on a car’s dashboard – of an important area of a judiciary’s performance.
Relevant: links to what matters, and counts what counts, in a meaningful and worthwhile way (i.e., consistent with visions, values, and strategic goals).

Time-bound: includes a time-frame for accomplishment (e.g., six months or one year).

The final selection of measures needs to meet these SMART criteria. They are not met by any of the ENCJ indicators in their current iteration. Most are merely a restatement of the aspects, elements, or issues of judicial independence and accountability. Similarly, in many cases, the sub-indicators are little more than a restatement of the indicators (e.g., code of judicial ethics, an indicator, is made into the sub-indicator, availability of code of ethics, with only the words “existence of” or “availability” added).

4.4. Management, not just measurement

**ENCJ and its Members should aspire to manage judicial independence, accountability, and transparency, not just measure performance in these areas. Paying attention to the management of performance -- such as the delivery of the “right” indicators to the “right” people, at the “right” time and the “right way” -- will pay great dividends even in the early stages of the development of indicators and measures.**

Successful leaders and managers see performance measures and indicators as powerful tools to communicate strategy and to change the behavior of individuals and groups. The best performance measures are actionable. Focusing on the management of performance, even in the early stages of identifying the right measures, by ensuring that identified measures are actionable and will lead to better performance measurement as well as better performance management. A surefire way of undermining the building of a successful PMM system is to create measures that are too abstract and poorly defined. Users of measures cannot improve performance if they do not understand what a measure means and what specific steps they may need to take to make use of the measure.

4.4.1. Key requirements and phases of development of indicators

Developing the right performance measures for an individual justice institution or an entire country’s justice system, and making sure that they are used effectively, can be translated operationally into three overlapping and interdependent key requirements and corresponding phases of development outlined below. These three requirements and phases of development are described in sequential fashion. In practice, they are iterative. The third requirement and phase, the “right” use, for example, is likely to need addressing early in the development process of performance measures to ensure that judicial leaders and top management are engaged early and remain so throughout the development of the PMM. While it may be frustrating for the developers, it is inevitable that the early question of the “right” measures will need to be answered in the context of satisfactory answers to the questions of the right use.

- **Right Measures** - Identifying and developing the right performance measures is much more than simply identifying and naming them. They must be painstakingly designed and developed. As described in preceding section, the right measures are SMART -- specific, measurable, attainable, relevant, and time-bound.
- **Right Delivery and Distribution of Performance Data** - Ensuring that the right measures are delivered to the right people, at the right time, and in the right way. Increasingly, this is done by information technology -- including performance dashboards, business intelligence, and data visualization applications -- that let users view critical performance information at a glance, and move easily through successive layers of strategic, tactical and operational information on a self-help, on-demand basis, allowing them to spot patterns, anomalies, proportions, and relationships that they would otherwise miss.53
- **Right Use** - Adopting, implementing, and integrating the measures of performance, as well as the delivery system and distribution system (e.g., performance “scorecards” and “dashboards”), with key management processes and operations, including budgeting and finance, resource and workload allocation, strategic planning, organizational management, and staff development.

4.4.2. Numbers should support but not replace judgement

The performance measurement process uses numbers to provide an understandable and comparative result, but it is ultimately not about the numbers. It is about perception, understanding, and insight to inform better decision making. Even strong advocates of numbers and mathematical computations aver that numbers should support but never replace our judgement.54 Ultimately, it is not the measure itself that is important, but rather the questions outlined below that it compels us to confront. This part of the recommendation may appear nonsensical and at odds with the essence of PMM. It is not. If you are asking the wrong questions or don’t act in response to the right questions, the answers really don’t matter.


The preceding sections explored risks to judicial independence and accountability at two levels: decisional independence (e.g., when a judge perfunctorily places a court’s imprimatur on decisions made by the other branches of government) and institutional independence (e.g., when the executive branch meddles improperly in judicial selection, retentions, and assignments). Today, there are transnational developments of risks at a third level – governance independence -- that are more troubling than risks to decisional independence and institutional independence. Some see these developments as a retreat from democracy and the rule of law and a movement toward “illiberal democracy” around the world that pose an existential threat to the judiciary as a coequal partner in government? In many countries, lip service is paid to democratic principles in constitutions that are ignored in practice.

Strongmen who rule in authoritarian Russian-style oligarchies amid corruption and political chaos in countries that once looked like they were democrats today seem to be moving in opposite directions. This is occurring not only in Russia, Egypt, and in Turkey, but also in my country, the United States, long thought to be the engine room of democracy. As they cling to power, embattled strongmen elsewhere, like Daniel Ortega in Nicaragua and Nicholas Maduro in Venezuela, show no inclination towards sharing power with others, never mind their own judiciaries.

Closer to home for the councils of the ENCJ is Hungary’s Prime Minister Viktor Orban who touts the virtues of illiberal democracy and who, no doubt, is willing to dispense with such constitutional niceties as an independent judiciary and the separation of powers because, he might argue, judicial independence impedes economic growth by making it harder for his government to respond quickly and flexibly to changing circumstances and national crises. Other leaders in Poland, the Czech Republic, and Austria, and further afield in Eastern Europe in the Ukraine and Moldova, all seem to be following Orban’s playbook.

Individual judges are, of course, constrained in their positions as judicial officers to respond to such existential threats to democracy. Entering the political arena in their judicial capacity risks violating the very principles they pledged to uphold. However, this is an area where ENCJ and individual councils can do much good. For example, they can assume an invaluable role by leveraging ENJC’s influence as a multilateral organization and by, for example, referencing treaties that promote judicial independence and the separation of powers. ENJC and its partners can give voice and agency to judges in support of judicial independence, separation of powers, and the rule of law in their roles and capacities as representatives of councils of the judiciary that otherwise might not be possible.

ENCJ and national councils could take several approaches to advocacy of judicial independence that judges in their official capacity as judges cannot easily take. Such efforts can establish norms and set up obstacles to authoritarian forces that flaunt judicial independence. For example, they could urge their countries to redouble commitments to be subject to the jurisdiction of international tribunals like the International Criminal Court at the Hague. They could encourage international cooperation and exchange of

55 Research by Professors Hayo and Voigt in 2005, cited by van Dijk and Vos, in their lead article, page citation shows that measures of independence are dependent on the extent of democratization. This may be the proverbial elephant in the room that cannot be ignored as we consider judicial independence and accountability.

information through various international treaties that promise to curtail strongmen in their exercise of power at the expense of judicial independence. Finally, their support of emerging technologies like blockchain and artificial intelligence that decentralize fundamental rights away from autocratic politician also have the potential to protect judicial independence.57

These approaches can be anchored in PMM, specifically ENCJ’s development of indicators of judicial independence and judicial accountability. PMM is not just a grand diagnostic or accounting exercise but also an instrument of power and control. There is today no longer much doubt whether good measures and indicators of performance can help achieve worthy ends. The question is whose values and vision of justice and court excellence they advance.58 Being able to “control events” (in the words of Sir Francis Bacon quoted at the beginning of this Conclusion), by determining the single version of the “truth” about judicial independence and accountability, by determining patterns and trends, and by measuring successes and failures, is as much an exertion of authority, power, and control, as it is a diagnostic exercise.


Conceptualization(s) of Judicial Independence and Judicial Accountability by the European Network of Councils for the Judiciary: Two Steps Forward, One Step Back

David Kosař and Samuel Spáč*

Abstract

This article focuses on conceptual issues regarding the new methodology of the European Network of Councils for the Judiciary (ENCJ) for measuring judicial independence and accountability. First, we argue that the proposal mixes up several concepts – judicial independence, judicial accountability, transparency of the judiciary, and public trust in the judiciary – which should be treated separately. Second, the proposal relies too much on conceptions of independence developed by the judicial community. As a result, it treats judicial administration with higher levels of involvement of judges as inherently better without empirical evidence, and does not sufficiently distinguish between de iure and de facto judicial independence. Moreover, the ENCJ’s indicators of judicial accountability are underinclusive as well as overinclusive and do not correspond to the traditional understanding of the concept. Finally, we argue that the ENCJ has to accept the possibility that (at least some types of) judicial councils (at least in some jurisdictions) might negatively affect (at least some facets of) judicial independence and judicial accountability. As a result, the ENCJ must adjust the relevant indicators accordingly.

Keywords: European Network of Councils for the Judiciary, Judicial Independence and Accountability, Measurement, Indicators

1. Introduction

The European Network of Councils for the Judiciary (hereinafter “ENCJ”) launched its ambitious project to develop indicators for measuring the independence and accountability of the European judicial system in 2013. It published its first report in 2014 and soon initiated its survey among judges to study perceptions of judicial independence. Then it fine-tuned the performance indicators of independence and accountability, and managed to received responses from 11,712 judges across Europe in its 2016-2017 edition of the survey. During those years the ENCJ worked hard to improve its understanding of judicial independence and accountability and, in doing so, greatly advanced our knowledge of European judicial systems.

This short article zeroes in exclusively on the conceptual issues regarding the ENCJ’s methodology, as developed by van Dijk and Vos in their leading article to this special issue. More specifically, it focuses on the ENCJ’s conceptualization of judicial independence and accountability, and then makes a few suggestions about how to make the ENCJ’s reports even more valuable for policy makers, scholars and other stakeholders. First, we argue that the current ENCJ report mixes up together too many concepts – judicial independence, judicial accountability, transparency of the judiciary, and public trust in the judiciary – that should be treated separately. Second, the ENCJ report does not sufficiently distinguish between de iure and de facto judicial independence. Moreover, it does not give proper credit to well-known experiences in some juridical systems where judicial self-government proved to be perilous to the judiciary. These experiences demonstrate the necessity to widen the focus to a variety of informal mechanisms, which can be easily misused. If we overlook these misuses, we might get a distorted picture of the functioning of the judiciary. Third, the ENCJ’s conceptualization of judicial accountability does not correspond to the traditional understanding of the concept. The ENCJ’s definition of accountability is underinclusive as well as overinclusive. More specifically, it overlooks well-known accountability mechanisms (such as disciplining of judges), while at the same time it includes mechanisms (such as understandable procedures and recusal) that have little to do with holding judges to account.

We sincerely believe that if the ENCJ manages to fix these conceptual issues, it might avoid several problems in the operationalization of performance indicators and in the subsequent steps in its analysis of judicial independence and accountability addressed by other articles in this special issue.

* David Kosař is associate professor, Samuel Spáč* is specialist at The judicial Studies institute of Masaryk University, Brno, Czech Republic. E-mail: David.Kosar@law.muni.cz

1 We therefore leave aside, among other things, the issues of operationalization, coding, aggregation of data, and reliability of the used sources.


2. Judicial Independence: connecting means and ends

In this part we focus on two broader conceptual issues of van Dijk and Vos’s proposal to measure judicial independence that has been utilized by the ENCJ in its surveys. First, the methodology relies too heavily in its definition of judicial independence on standards developed by the judicial community and does not pay sufficient attention to well-documented threats to judicial independence coming from within the judiciary. Second, the operationalization of judicial independence conflates the approach focused on independence as a feature of institutional design and the approach focused on judicial decision-making. Moreover, van Dijk and Vos rely on data measuring public trust in the judiciary instead of judicial independence, which is the central concept of the proposed study.

2.1. Conceptual consequences of reliance on international standards

The conceptualization of judicial independence proposed by van Dijk and Vos relies too heavily on the standards developed, to a large extent, by the international judicial community. Consequently, it leads to a problematic assumption present throughout the methodology that judicial independence is best served when judicial administration is carried out predominantly by judges appealing to *communis opinio*, as international documents that provided some recommendations on how to secure judicial independence really hold a similar position. However, the usefulness of this approach has been criticized on numerous occasions.

With regard to the proposed methodology and its future application in (and beyond) Europe it is problematic from at least two perspectives. First, it largely ignores a literature that shows that threats to judicial conduct may come from inside the judiciary as well. Second, since the conceptualization of judicial accountability in the proposed methodological design stresses the importance of answerability of the judiciary to the general public, it is surprising that as regards independence the influence of democratically elected politicians creating a link of legitimacy between the judiciary and the demos is portrayed as a danger to “the balance of state powers”.

Judicial independence is, first and foremost, a relational concept. It focuses on connections between the judiciary and other actors in the political arena – most usually on the interaction with the political branches of power. In this vein, Van Dijk and Vos argue that independence is secured when the judicial power is protected from political branches, for instance by arguing that “it is only logical that, for the judiciary to play its constitutional role in the checks and balances between the state powers in a democracy, it should not be under the control of other state powers.”

This argument is certainly valid, but the checks and balances do not operate in only one direction. More specifically, the courts are not just checks upon other branches, but are also subject to checks by those branches. The same applies to the balances that also operate in both ways. Other approaches posit that judicial independence ought to secure that judges are free to decide cases according to law and evidence, in order to secure equality before law. To be fair, van Dijk and Vos also present this argument. However, the indicators in their methodology focus primarily on de jure aspects of the issue with considerable preference for systems where judicial councils dominated by judges represent the independence of the judiciary.

We argue in favor of analyzing judicial independence, and its presence or absence in any polity, with a focus on undue influence rather than separation of the judiciary from political branches. The institutional approach fails to give due credit to the empirical evidence showing that judicial conduct can be threatened from inside as well as from outside of the judiciary. That means that the ENCJ needs to study not...
only external, but also internal independence. In fact, even the ECHR recognizes the requirement of internal independence of judges and the dangers of judges being influenced by their colleagues and superiors.

As several studies have shown, judicial self-government may lead to misuse of disciplinary procedures and other accountability mechanisms, distortion of merit-based selection system, or can cause encapsulation of the judiciary and its disappearance from the public eye. To put it bluntly, there is no particular reason (at least in some societies) that judges will behave any differently than politicians when using these powers. At the same time, there are several examples suggesting that when politicians have formal powers to interfere with judicial careers, they may to a large extent opt out from using them.

Consequently, it is necessary to take into consideration that there always will be actors empowered to influence the judiciary, and that there is no reason to believe that judges, as a group or its individual leaders, do not have any selfish interests just like any other actors. Therefore, judicial independence should rather be understood as a ‘consequence of self-restraint by powerful actors’, and less as the feature of a particular institutional design. Certainly, institutional design can make it more difficult for powerful actors to exert influence over the judiciary, but, to put it simply, it cannot preclude the existence of ‘powerful actors’ and it cannot prevent them from influencing courts altogether. As a consequence, it is necessary to distinguish between the capacity of these powerful actors – understood as existing channels to exercise pressure on courts, and their willingness – a result of conscious choice, to utilize this capacity.

There are several reasons why powerful actors may not be willing to pressure courts and exhibit self-restraint. It can be the result of their strategic calculation suggesting that benefits do not outweigh costs, or it can be their honest adherence to values of democracy and the rule of law. In addition, the judiciary itself can pose as a counterweight to powerful actors’ whimsy, and certainly can resist the attempts of undue influence – be it on the basis of their own belief in the rule of law, their understanding of their role in the broader context, cost-benefit analysis, or perhaps it can depend on their social legitimacy or support from the media.

To put it differently, proposed conceptualization and operationalization of judicial independence seem to be skewed in favor of countries that followed a variety of international recommendations and transferred considerable powers into the hands of the judiciary. However, as several scholars have argued before and as empirical evidence suggests, the transfer of formal powers from political branches to the judiciary does not necessarily decrease the threat of misuse of these powers, but rather should expand our attention to other actors (including those within the judiciary) as well. Relying on assumptions which hold that the judiciary governed by judges is an independent judiciary treats judicial independence, in institutional terms, as an end in itself, even if it does not bring desired outcomes.

### 2.2. Reconciling de iure and de facto definitions

Probably all conceptualizations of judicial independence fall in the spectrum between perceiving independence as a feature of institutional design, and as a feature of judicial decision-making. Various terms can be applied to this dichotomy: some differ between independence as means and ends; others apply labels such as ‘structural insulation’ or refer to judicial independence as a specific mechanism.

---

16 Bojarski and Stemker Köster, supra note 13; Kosař, supra note 7, pp. 317-329.
17 Spáč, supra note 7.
22 Popova, supra note 12, pp. 20-23.
27 Popova, supra note 12.
when speaking about the former; while others perceive it as a value in itself, or label it as ‘party detachment’, “impartiality” or ‘judicial autonomy’, when speaking about the latter. One can also distinguish between institutional, decisional and behavioral independence, here the first term focuses on the level of the judiciary, while the other are understood at the level of individual judges.

The methodology proposed by van Dijk and Vos correctly distinguishes between the two approaches by differentiating between independence on the level of the judiciary as a whole and on the level of individual judges. However, authors do in this regard make at least three noteworthy methodological choices. First, they do not treat the two approaches separately, even though they measure different (although possibly related) dimensions of the concept. Second, they overlook an important link between the two, leaving them rather unconnected (yet analyzed as a single concept) without sufficient attention being paid to how institutional design translates into courts’ ability to deliver independent outcomes. Finally, to assess independence in terms of judicial decision-making the authors use perception indicators, considerably resembling indicators for a completely different concept – public trust.

As regards the relationship between independence understood in institutional terms and independence on the level of judicial decision-making, several authors have addressed the issue. According to Tiede, for a judiciary to be considered independent, it first needs to be independent institutionally, and only afterwards can its independence be measured as the amount of discretion judges enjoy. Russell, like van Dijk and Vos, distinguishes between collective and individual targets of undue influence, while arguing that a greater amount of institutional autonomy does not secure ‘that judges will think and act in an independent manner.’ However, if they do not enjoy autonomy in institutional terms, odds on their defiance to powerful actors decrease. Burbank and Friedman in a comparable manner suggest that independence should be perceived as a means to an end, not an end in itself. In general, there seems to be an assumed link between independence in institutional terms and independence understood at the level of individual judges. Nevertheless, the former should not be perceived as a sine qua non for the latter. Even comparative empirical evidence fails to support the existence of a strong relationship between the two. All in all, the literature suggests the two approaches should be treated separately as they may or may not be associated.

We argue that the missing link between the two approaches is the way in which powerful actors utilize their formal powers to modify the composition of the judiciary. As was shown by previous research, de facto judicial independence, when operationalized as the utilization of formal powers, may serve as a predictor for variables related to well-performing judicial systems. According to Popova, an independent judiciary produces decisions that do not show a consistent preference for any actor or group of actors. There are two reasons why researchers should be concerned with the way formal powers translate into the composition of the judiciary. First, powerful actors can systematically favor judges (either consciously or subconsciously) with certain characteristics, e.g. in the process of the selection or promotion of judges, in such a way that no other pressure would be necessary to secure desired outcomes. However, skewing the composition of the bench through selection and promotion of judges is a rather long-term project requiring concentrated action for a longer period. Contrarily, if powerful actors use their powers to discipline, transfer or dismiss judges they can effectively demonstrate their capacities to the judiciary motivating judges to certain actions, or discouraging them from others, in a shorter time period. In sum, we hold that to study judicial independence one should look at three levels: distribution of formal powers, utilization of these powers to alter the judiciary, and the output of judicial conduct, while treating them separately in order to understand how they are interconnected.

Finally, as regards the proposed indicators of the subjective independence of the judiciary and individual judges, we understand authors’ choice to use data from opinion surveys as a proxy. However, it needs to be borne in mind that more than what they want to measure (judicial independence) they may in fact measure public trust, a concept not necessarily related to independence. Public trust/confidence in the judiciary can be understood as the public’s belief in the reliability and honesty of courts and judges, and their ability to competently

---

33 Popova, supra note 12, pp. 14-19.
34 Tiede, supra note 11, pp. 133-135.
35 Russell, supra note 9, p. 7.
36 Burbank and Friedman, supra note 26.
38 Voigt et al., supra note 6.
39 We do understand that to study output of the judicial conduct in all European judiciaries is beyond the capabilities of the ENCJ. However, the ENCJ could and should distinguish between distribution of formal powers (de iure independence) and utilization of these powers to alter the judiciary (de facto independence).
perform functions that are assigned to them.\textsuperscript{41} To be sure, perception of independence may be a substantial part of such an evaluation, but not necessarily the most important one.

3. Judicial accountability: in search of a viable conceptualization of an understudied concept\textsuperscript{42}

In this part we argue that the conceptualization of judicial accountability used by the ENJC reports and defended by van Dijk and Vos in their leading article is even more problematic than ENJC's conceptualization of judicial independence. Most importantly, this part of the ENJC's assessment does not communicate with the existing (judicial) accountability literature at all,\textsuperscript{43} which results in several conceptual problems. First, while the ENJC report does not provide any definition of accountability, the definition proposed by van Dijk and Vos is tautological, vague and normatively loaded, which makes it unsuitable for comparative analysis. Moreover, merging accountability with transparency causes serious methodological concerns. Second, we show that several indicators measured under the ENJC's judicial accountability assessment have little to do with holding judges to account in the traditional sense, while at the same time the ENJC overlooks well-known accountability mechanisms such as disciplinary measures. Third, we propose a means of moving forward with the conceptualization of judicial accountability.

3.1. We need a definition of judicial accountability before moving to indicators

The concept of accountability is notoriously difficult to define. The very fact that this term lacks a proper equivalent in most non-English languages makes the inquiry difficult. What is more, lawyers lag behind political scientists\textsuperscript{44} and economists\textsuperscript{45} in analyzing this concept. As one of us lamented a few years ago when reviewing three books on judicial accountability, ‘none of the authors devotes much attention to the conceptual question: what exactly is meant by accountability?’ and that ‘[t]his deficit, in turn, results in conceptual confusion that hampers scholarly progress’.\textsuperscript{46}

However, since then the relevant political science and economics literature has become widely known among lawyers.\textsuperscript{47} As a result, it became clear that in order to save the term judicial accountability from becoming a meaningless ‘buzz-word’ and to make it analytically useful, every study on judicial accountability should take a stance on two issues. First, it should make clear whether it is studying judicial accountability as a virtue or as a mechanism.\textsuperscript{48} Secondly, it should signal the scope of its inquiry to the reader by providing answers to three fundamental questions, namely: (1) is the term ‘mechanisms of judicial accountability’ reserved solely for ex post mechanisms?; (2) must accountability mechanisms entail consequences\textsuperscript{49} for the actor that is held to account?; and (3) does accountability encompass only sanctions or both sanctions and rewards?

In a recent study on the impact of introducing into Slovakia the judicial council on judicial accountability, one of us made clear that he studies accountability as a mechanism, and that he defines accountability mechanisms as ex post mechanisms that must entail consequences, in the form either of a sanction or a reward. This resulted in the following definition of judicial accountability: ‘a negative or positive consequence that an individual judge expects to face from one or more principals (from the executive and/or from the legislature and/or from the court presidents and/or from other actors) in the event that his behavior and/or decisions deviate too much from a generally positive consequence that an individual judge expects to face from one or more principals (from the executive and/or from the legislature and/or from the court presidents and/or from other actors).’

That stems from Kosař’s particular definition of the adjective “judicial”, which we cannot discuss here in more detail.


\textsuperscript{42} CCJE, Opinion No. 18 (2015) to be taken into account.


\textsuperscript{44} Bovens, supra note 43; Schedler, supra note 43; Ferejohn, supra note 43.

\textsuperscript{45} Voigt, supra note 43.

\textsuperscript{46} Kosař, supra note 43.


\textsuperscript{48} See Bovens, supra note 43.

\textsuperscript{49} In other words, must the principal be able to impose sanctions or grant rewards?

\textsuperscript{50} Kosař, supra note 7.

\textsuperscript{51} That stems from Kosař’s particular definition of the adjective “judicial”, which we cannot discuss here in more detail.
Secondly, merging accountability and transparency into one concept blurs rather than clarifies things, as transparency is in itself a complex correlation should be distinguished from causality.

Moreover, their definition restricts accountability only to accountability to the society without any justification. Why is accountability to the legislature, which represents the people, ruled out? In many developed European societies, the democratic principle forms the backbone of the constitution and may sometimes even be protected by the unamendable part of the constitution, the so-called eternity clause.65 Again, van Dijk and Vos’s definition is normatively loaded, as it implicitly finds accountability to the elected politicians inappropriate and thus outside their definition.

Yet another conceptual problem arises from the conflation of accountability and transparency. More specifically, van Dijk and Vos claim that ‘[a]ccountability and transparency are … closely linked’ and thus they ‘use these terms interchangeably’.57 While there are some broad conceptions of accountability that take a similar approach,58 such merger suffers from two problems.

First, it means that judicial accountability is defined without the ‘enforcement’ pillar69 and implies that if the judiciary is transparent (and under public control), it will act properly. That is a very bold assumption. Has transparency ever changed performance in itself? More specifically, why and how should a transparent judiciary change performance if there is no one particularly responsible to the public? What if the courts were forced to be transparent by an external actor, such as the parliament that found the functioning of the judiciary opaque and problematic?60 While transparency may very well contribute to the change in judicial performance, it is not necessarily so and the mere correlation should be distinguished from causality.

Secondly, merging accountability and transparency into one concept blurs rather than clarifies things, as transparency is in itself a complex concept. Transparency is not only about accessibility, but also about findability and understandability of the accessible information.61 Information about courts and judges is available when it is possible to access it. Information about courts and judges is findable if it is visible – it reflects the degree to which information can easily be located. For instance, if information is available in the open access regime (rather than upon request or even subject to fees) and via many channels (online, in press releases, in a specialized journal etc.), it is more findable. Finally, understandability of information reflects the extent to which it can be used to draw accurate conclusions. For instance, if information is provided in the user-friendly format62 (e.g. in aggregated form, as a summary or commented document rather than just raw data that require time-consuming and resource-heavy processing), the reader can easily infer the findings and their significance, and thus such information is more understandable. Merging all of this under the heading of accountability makes the concept of accountability extremely complex.

This is not to say that the transparency of the judiciary should not be studied by the ENCJ. To the contrary, it is an important value. We also agree that transparency and accountability are interrelated, at least in the sense that transparency is a prerequisite for accountability, but they should be treated separately. In sum, transparency is a separate concept that operates rather as the contingent circumstance that might influence whether a certain form of accountability will bring about a certain set of results.63

52 See van Dijk & Vos, supra note 2, p.7.
53 See Voigt, supra note 3.
54 The same problem arises, mutatis mutandis, if we treat the “rule of law” as the “rule of good law”. See J. Raz, The Authority of Law: Essays on Law and Morality (OUP 1979).
55 This is a standard position of all relevant stakeholders in Austria.
57 See van Dijk & Vos, supra note 2, p. TBA.
58 Bovens, supra note 43.
59 Schedler, supra note 43.
62 n fact, the following three attributes increase understandability (inferability) of data: disaggregation (raw data are the least biased), verifiability (ability to be marked as right data by a third party), and simplification (providing tools for lay people to understand the data, codebooks etc.). See Michener and Bersch, supra note 61.
63 For a similar approach, see M. Philp, Delimiting Democratic Accountability, Political Studies 57, pp. 28-53.
3.2. Accountability indicators: overinclusive as well as underinclusive

The abovementioned conceptual problems translate into a very peculiar selection of accountability indicators and sub-indicators that is far away from a traditional understanding of accountability mechanisms. More specifically, the ENCJ’s set of indicators is both overinclusive and underinclusive.

Out of the indicators of the objective accountability of the judiciary as a whole, only the complaints procedure and judicial performance evaluation (periodic and public benchmarking of the courts) are accountability mechanisms under any conception of accountability as they entail consequences for the judiciary. It is less clear how the mere publication of the annual reports, the transparent allocation of cases, and relations with the press in themselves may hold the judiciary to account. These three measures tell us more about transparency than about accountability, since without further action they do not result in any consequences for the judiciary. The indicator of “external review” is then too vague as it is not clear what it means if the judiciary is open to external review.

The indicators of the objective accountability of individual judges are even more problematic. A code of ethics in itself does not hold individual judges to account. In several countries, judges openly ignore the code of ethics and do not face any consequences. It is only the subsequent disciplinary action that holds the given judge who breached the code of ethics accountable. The same applies to recusal, admissibility of external functions, and disclosure of financial interests. These indicators just set the standard, but do not ensure that someone will enforce it against recalcitrant judges. Finally, understandable proceedings are totally unrelated to any plausible conception of judicial accountability.

At the same time, the set of indicators in van Dijk and Vos’s article is underinclusive. Typical accountability mechanisms of the judiciary include budgeting for the courts, reducing the compulsory retirement age of judges, and the dismissal of court presidents. As to the accountability of individual judges, one of us has provided a detailed account of this aspect of judicial accountability elsewhere, and thus we will not elaborate on it here. It suffices to say that virtually any definition of accountability of individual judges includes disciplinary proceedings. It is unclear to us why both the ENCJ and van Dijk and Vos’s article exclude this mechanism from their account of accountability. This of course is not to suggest that disciplinary proceedings are the only accountability mechanism. The other mechanisms may include, among other things, impeachment, retention, relocation, reassignment, and demotion of judges, volatile judicial salaries, case assignment and reassignment. Broader conceptions of accountability may also accept the election of judges as a method by which judges are held to account.

3.3. How to move forward?

In order to improve our knowledge of how judges are held to account in different societies, we urge the ENCJ to sharpen its definition of judicial accountability by avoiding the current vague moralizing language, strip it from the normatively indefensible limitations (such as a very narrow understanding of the accountability of whom question), and to distinguish it clearly from adjacent concepts such as judicial transparency and from wholly different issues such as recusal of judges and understandable proceedings.

Irrespective of how broad or narrow a definition of accountability the ENCJ eventually adopts, we believe that the ENCJ should focus on three dimensions of judicial accountability: (1) de iure accountability of both the judiciary and individual judges; (2) de facto accountability of both the judiciary and individual judges, and (3) informal accountability of both the judiciary and individual judges. De iure judicial accountability should ask who has the formal power (as defined by legal rules) to hold the judiciary and judges to account and for what. De facto judicial accountability should probe into how these formal rules operate in practice and which formally empowered actors actually hold the judiciary and judges to account in practice and for what. Finally, informal judicial accountability would explore whether there are any informal mechanisms (not regulated by law) via which the judiciary and judges are held to account. Such informal mechanisms may include, among other things, forced retirements, media pressure, powers of court presidents to assign subsidized flats and to decide on attending foreign conferences, and the ruling party implicating a judge’s family in a scandal.

In the long term, the ENCJ might also advance our understanding of perceptions of judicial accountability (i.e. subjective judicial accountability in the ENCJ’s parlance). To start with, the ENCJ could add a new question to the ENCJ survey that would ask European judges about “accountability as perceived by judges”. In addition, it may initiate the inclusion of a similar question to the Flash Eurobarometer, to national surveys among court users, and to the EU Anti-Corruption reports.

64 See e.g. Iancu, supra note 18 (regarding Romania); Kosař, supra note 7, pp. 317-329; S. Spáč et al., supra note 60; and Spáč, supra note 7; and (regarding Slovakia).
65 Kosař, supra note 7.
66 A. Le Sueur, supra note 43; Kosař, supra note 7; Bovens, supra note 43; Voigt, supra note 43; D. Piana, supra note 43.
67 One possible explanation would be that the ENCJ treats judicial independence and judicial accountability as mutually exclusive concepts. However, this is an indefensible position (even van Dijk and Vos themselves in their article acknowledge that certain accountability indicators such as allocation of cases also raise independence concerns, see van Dijk and Vos, pp. TBA).
68 Kosař, supra note 7.
69 On the distinction between de iure and de facto judicial accountability, see Voigt, supra note 45; and Kosař, supra note 7 (especially chapters 1 and 2).
70 Ideally, the ENCJ should also inquire into through what processes are judges held to account, by what standards, and with what effects, but we have to simplify here. For more details, see J. Mashaw, Accountability and Institutional Design: Some Thoughts on the Grammar of Governance, in: M.W. Dowdle, Public Accountability: Designs, Dilemmas and Experiences (CUP 2006), p. 118; and Kosař, supra note 7, pp. 40-58.
4. Conclusion

We are extremely grateful that we were allowed to reflect on the ENCJ’s methodology and engage with van Dijk and Vos’s article. It should be particularly appreciated that the ENCJ opened this debate and invited practitioners and scholars from different fields and with different views on the subject matter. It is a bold move that only a few international associations are willing to make. We did our best to contribute to the development of the ENCJ’s conceptualization of judicial independence and accountability. We also understand that one of the key objectives of the ENCJ’s Project on Independence and Accountability is ‘to present an ENCJ vision on the independence and accountability of the judiciary’. This vision is based on the assumption that judicial independence is best served when judicial administration is carried out predominantly by judges. However, this approach, even though a part of the raison d’être of the ENCJ, at best prioritizes judicial supremacy despite unclear evidence about real performance of judicial self-government, and at worst overlooks the growing evidence that judicial councils may actually weaken judicial independence.

Therefore, there are some hard decisions to be made by the ENCJ. Most importantly, it needs to accept the possibility that (at least some types of) judicial councils (at least in some jurisdictions) might negatively affect (at least some facets of) judicial independence and judicial accountability. If the ENCJ contributes to our knowledge of how this is done (by what means and which actors), it would in itself be a great achievement. We believe that the ENCJ is particularly suited to this task, as judges know best what negatively affects their independence and accountability. And if the ENCJ manages to finetune its indicators so that the revised operationalization allows us even to identify which organizational traits of judicial councils matter and under what conditions judicial councils increase both judicial independence and judicial accountability (and, potentially, also the transparency of the judiciary and public trust in courts), it would be a tremendous breakthrough.

Bibliography

Bochan v. Ukraine, 7577/02 (European Court of Human Rights May 3, 2007).
Daktaras v. Lithuania, 42095/98 (European Court of Human Rights October 10, 2000).
DMD GROUP, a.s. v. Slovakia, 19334/03 (European Court of Human Rights October 5, 2010).

---

72 ENCJ, ENCJ project on Independence and Accountability, available at: https://www.encj.eu/articles/71 (emphasis added).


Gazeta Ukraina-Tsentr v. Ukraine, 16695/04 (European Court of Human Rights July 15, 2010).


Miracle Europe Kft. v. Hungary, 57774/13 (European Court of Human Rights January 12, 2016).

Moiseyev v. Russia, 62936/00 (European Court of Human Rights October 9, 2008).


Sutyagin v. Russia, 30024/02 (European Court of Human Rights May 3, 2011).


Innovate – Don’t Imitate!
- ENCJ Research Should Focus on Research Gaps

Stefan Voigt*

Abstract:
After a critical evaluation of the proposal by van Dijk and Vos to make judicial independence and accountability measurable, this paper discusses a number of basic methodological decisions that need to be made before meaningful indicators can be constructed. The paper then moves on to deplore the lack of research regarding the effects of judicial councils and proposes that more research be carried out to reduce our knowledge deficit. It further describes the rising importance of prosecutors and the complete lack of indicators regarding prosecutorial accountability. It argues that producing indicators with regard to prosecutors could add a lot more value than yet another indicator on judicial independence.

Key words: judicial independence, judicial accountability, prosecutorial independence, prosecutorial accountability, judicial councils, indicator construction.

Innovate – Don’t Imitate! – ENCJ Research should focus on Research Gaps

1 Introduction
Reliable measures for both judicial independence (JI) and judicial accountability (JA) are a desideratum. Only if we are able to reliably measure the degree to which various countries have realized the independence and accountability of their judiciaries are we able to ascertain their effects empirically. Policy recommendations aiming to improve JI and JA will only be convincing if the precise weaknesses that a judiciary suffers from are well known. Those are only two important reasons for why reliable measures are dearly needed. In principle, then, the attempt of the ENCJ to help establish such indicators is highly welcome and the responsible ENCJ bodies are to be congratulated for their initiative.

But there is an important limitation. The proposal by van Dijk and Vos to be discussed here does not really get us closer to reliable indicators. Let me briefly summarize. The proposal rests on questionable assumptions, the most important of which is that the judicial councils are conducive to both JI and JA. Empirical evidence in support of this assumption is lacking and relevant studies seem to indicate the exact opposite. Furthermore, the proposed indicators are undertheorized, some variables included in the proposed indicator are fraught with problems and the proposed way to aggregate the single variables into an overall indicator is flawed. Several indicators that suffer neither from the theoretical nor the empirical flaws of this proposal are readily available. In other words: It is not clear what the new indicators could contribute to the debate on how to optimally design the judiciary.

How to get out of this mess? I suggest to rely on existing indicators and redefine this data gathering project. The ENCJ should focus on its strengths. It has the capacity to get in touch with thousands of judges – and possibly also with thousands of court users. There are many things these people know that are not commonly known – and are of relevance to the perception as well as the functioning of judicial systems. Tapping this resource could add significant value to the academic and policy debate.

The remainder of my contribution is structured in the following way: In the next section, I explicitly deal with some of the flaws of the proposal by van Dijk and Vos. The following section discusses a number of methodological decisions that need to be taken before constructing an indicator. In a sense, it serves to offer some background regarding the heavy criticism in Section 2. In Section 4, I encourage the ENCJ bodies to be innovative and sketch one or two projects where a real value added seems possible.

2 Shortcomings of the Proposal
In the introduction to their paper, the authors write that “the lack of a comprehensive approach in the assessment of judicial independence and judicial accountability is striking.” They conclude that “most studies are of an ad-hoc nature”, somewhat implying that theirs is not. I think that both claims are massively exaggerated.
Synergy Case Management System (CMS) is an integrated system for managing the entire court case lifecycle from initial filing through disposition. Synergy CMS automates case processing within individual courts, and among justice sector institutions, with integrated modules available for police, prosecutors, courts, and corrections. Built-in analytical tools enable courts to track court statistics and monitor performance.

Synergy CMS currently processes over 20,000 cases monthly, serving thousands of justice-sector employees and countless public users.

Synergy’s solutions have been implemented in more than 65 countries throughout Africa, Australia, Asia and the Pacific, the Caribbean, Central and Latin America, Europe and the Middle East.

www.synisys.com
Over the last 20 years, there has been a broad discussion in both academia and policy circles on how to assess JI and JA. Rick Messick’s (2001) paper on “Key Functions of Legal Systems with Suggested Performance Measure” was an important early contribution. The 2003 paper by Feld and Voigt contains objective (non-perception-based) measures for both *de jure* and *de facto* JI. These country indicators have not only led to many contributions on how the indicators could be improved but they have also been used in many empirical studies. In later years, a number of think tanks have invested ample resources to produce many more reliable indicators (many of which are included in the survey by Ríos-Figueroa and Staton 2014).

Compared to what is readily available, the current proposal is rather disappointing as it suffers from many flaws. Let me mention some of them. A necessary condition for producing reliable indicators is a precise definition of the key terms. A previous version of the paper did not contain any definition of the term “judicial accountability”. The authors have now added this sentence: “Accountability is used here in the sense of the judiciary being morally accountable to society in general.” This definition is puzzling for many reasons: It “defines” the noun accountability with the adjective accountable. By adding “morally” it becomes incomprehensible in what situation the defining elements of judicial accountability are present or absent. Moral intuitions are likely to differ across countries. Does that mean that the standard for JA should be context-dependent as well? If the authors are serious about this, then the attempt to come up with an indicator across countries is a non-starter. Moreover, accountability to the population is distinct from accountability to higher courts, to the law, to the law-makers, to disciplinary bodies, etc. In a judicial system that does not publicly elect judges, other forms of accountability appear to play a far more important role.

Some of the variables proposed to be included in the indicators are questionable, to say the least. On what theoretical grounds should there, e.g., be any “formal guarantees for involvement of judges in development of legal and judicial reform”? I might be old-fashioned, but the separation-of-powers doctrine would seem to say otherwise. Even if there were such guarantees, I do not see why that should necessarily lead to higher levels of JI.

Here is another example: the existence of a council for the judiciary is interpreted as one variable causing or reflecting high levels of JI. It is not only the existence of a council that counts, but also whether the council complies with EN CJ guidelines. True, councils have been high on the agenda of organizations like the World Bank who have encouraged many governments to adopt them. Yet, the empirical evidence regarding their effects that I am aware of sheds doubt on the assumption that they help making the judiciary more independent. In an early paper on judicial councils, Garoupa and Ginsburg (2009) find that neither the composition nor the competences of judicial councils are significant in explaining differences in the degree to which the rule of law is realized (of which JI is, of course, a central component). In a more recent paper, Voigt and El Bialy (2016) analyze the determinants of judicial performance based on data provided by the European Commission for the Efficiency of Justice (CEPEJ). To their surprise, judicial councils were consistently correlated with a worse, rather than a better performance. Although judicial efficiency is obviously not identical to JI, the sparse empirical evidence on the effects of judicial councils is sobering and raises the question whether promoting their introduction may have been ill-advised.

In aggregating single variables to an overall indicator, transparent and effective aggregation rules are needed. Attaching equal weight to all variables is common practice here and can be considered the default option. Attaching different weights to different variables makes sense, but needs to be thoroughly theorized. In the proposal under scrutiny here, the most frequently used scores are 0, 1, and 2. There are, however, exceptions: if it is not possible to transfer judges without their consent, the country receives 15 points. Why 15 and not any other number remains unclear. The authors do not provide any reasons for their coding decisions and the proposal is, hence, seriously undertheorized.¹

Probably the most serious and fatal flaw in the indicator’s construction is that countries for which no data on a specific variable is available, are simply coded 0. Two things are, hence, conflated here: country characteristics actually lowering JI on the one hand and lacking willingness or ability to provide data on the other. A country with a highly independent judiciary could, hence, score very badly simply because it did not report enough data.

From previous research, we do not only know that *de jure* and *de facto* JI are not very highly correlated but also that high levels of *de jure* JI have no positive (economic) effects, whereas *de facto* JI is highly correlated with economic growth (Feld and Voigt 2003, Voigt et al. 2015). This finding implies that reliable *de facto* indicators are of particular interest. Unfortunately, the current proposal does not contain such an indicator but relies, instead on “perceived independence.” In all likelihood perceived independence will be influenced by media reports and will thus be influenced by recent stories about cases in which the judiciary could not act independently and by the overall degree of press freedom. Regarding what is most likely the most important aspect of JI, namely its actually realized degree, the proposal remains essentially mute.

### 3. Methodological Considerations

Until now, my comment has been rather harsh and not very constructive. In the next two sections, I focus on suggestions for how the measurement project could be improved. This section contains a number of methodological considerations that should be dealt with before delving into the business of constructing indicators.

---

¹ Interestingly, the term “transfer” is nowhere specified and seems, hence, to be used in a very general fashion. It might have been interesting to separate different kinds of transfer, such as to different locations, to other types of courts, or other positions in the judicial system more generally (such as being transferred to become a prosecutor).
Choice of Method Depends on Underlying Goal

I assume that the underlying goal of the entire project is to help strengthen both JI and JA, e.g., because they are seen as important components of the rule of law or because they have positive economic effects (such as more transactions, more investment and, at the end of the day, faster economic growth). If this is the final goal, an intermediate goal would be to generate policy proposals conducive to that final goal. One precondition for getting there has already been named in the last section of this comment: the central terms need to be defined precisely. Whereas van Dijk and Vos start off by offering a definition of JI, JA is delineated in an extremely fuzzy way.

If one is indeed interested in the generation of policy proposals, it is important that indicators are “actionable.” Telling politicians merely that the JI score of their country is low will not be helpful in that regard. So, highlighting single aspects or dimensions that could be improved promises to be more helpful than focusing solely on one highly aggregated indicator. At the very least, it would have to be shown that sub-indicators tend to move in the same direction as the overall indicator.

That being said, a word of caution is in order: The low correlation between de jure and de facto JI shows that many countries have legislation conducive to JI which is, however, often not implemented on the ground. While incapacity might be one reason, the preferences of politicians might be another. When push comes to shove, many politicians seem to be ready to renge upon formal legislation and interfere into the business of a formally independent judiciary. Not taking the preferences of politicians into account while offering policy advice is therefore unlikely to have any beneficial effects.

But that is not the only reason to be cautious. Many programs aiming at improving the rule of law, of which JI and JA are important parts, miserably failed (e.g. Andrews 2013). And this outcome is likely even when we assume that politicians are genuinely interested in high levels of JI and JA because formal institutions – such as laws – can be radically changed overnight but it might take years or even decades until they really constrain and channel behavior in the way intended. Sounds like wild speculation? Here is some evidence: Based on data from the EU Justice Scoreboard, Gutmann and Voigt (2018a) show that national levels of judicial independence (as perceived by the citizens of EU member states) are negatively associated with the presence of formal legislation usually considered as conducive to judicial independence. The authors try to solve this puzzle based on political economy explanations and specificities of legal systems, but to no avail. They then ask whether cultural traits can help to put together the puzzle. And indeed, countries with high levels of generalized trust (and to a lesser extent individualist countries) exhibit increased levels of de facto judicial independence and, at the same time, reduced levels of de jure judicial independence. The combination of these two effects of culture on judicial institutions can explain why judicial reforms that should be conducive to an independent judiciary may seem to have adverse consequences. The authors conclude that cultural traits are of fundamental importance for the quality of formal institutions, even in societies as highly developed as the EU member states.

Objective vs. Subjective Indicators

In developing indicators, a decision needs to be taken whether one wants to rely on objective or subjective indicators (a mixture is, of course, possible, but not necessarily preferable). Subjective indicators are based on the perceptions of citizens or experts. The subjective indicators dealing with JI very frequently rely on assessments by so-called country experts or businesspeople. The Global Competitiveness Report of the World Economic Forum or the World Competitiveness Ranking by the IMD Business School are some of the best-known examples. They provide very interesting information. If business people perceive a country’s judiciary as not independent from government, they are likely to invest less. In that sense, they are not only interesting, but also highly relevant. They do, however, have an important drawback: these indicators are not actionable. If a government wants to increase its chances of receiving foreign direct investment, it needs to know the reasons why its judiciary is perceived as dependent.

This is a drawback from which objective indicators do not suffer. Since they are based on facts, they can be replicated by anybody interested in replicating them. This lends them higher credibility.

Here is a concrete proposal how the current project could add value to the academic and policy debate on how to organize the judiciary. I have argued that, at least in principle, both subjective and objective indicators can reveal relevant information. It would be helpful to identify those objectively verifiable factors that determine perceptions of JI. This could easily be turned into implementable policy advice. Few – if any – surveys have been carried out among court users. It would be great if the ENCJ could change this and inquire into the experiences of court users. Now, you have noticed that I am not talking of perceptions, but of experiences. Until now, many surveys include answers from people who have never used the courts themselves. They are, thus, revealing perceptions that might be formed on the basis of media reports, conversations with friends and family, and so forth. It would surely be interesting to run a survey solely with people who have first-hand experience with courts.2

4 Innovate – Don’t Duplicate!

The header to this section indicates that the extant project proposal duplicates existing indicators rather than adding something innovative. To substantiate that claim I begin this section by referring to three examples of readily available indicators regarding JI. In the second part of this section, I then sketch two projects with which the ENCJ working group could really add value.

2 Gutmann et al. (2015) have studied the relationship between perception vs. experience regarding corruption and find that they are not as highly correlated as one might expect and further that some basic sociodemographic characteristics can help to explain differences between the two.
Available Indicators

Over the last couple of decades, a very high number of academics and organizations have made indicators on JI available. There is great heterogeneity among them: some focus more on de jure than de facto, some draw on subjective perceptions whereas others try to code objective information. Since there are so many of them, I will not even try to name them all here but simply choose three as examples.

Linzer and Staton (2015) noticed that many of the available indicators were suffering from at least one flaw: they were only available for a single point in time (or rather few points in time) and they were only available for a rather small number of countries. Linzer and Staton offer a solution to this problem by aggregating in a sophisticated statistical measurement model the data from various sources to cover 200 countries over the time period since 1948. Their measurement model used for uncovering latent JI is able to address temporal dependence in observed and unobserved variables, substantial missing data, and measurement error in observed indicators, as well as conceptual boundedness in the latent quantity. Their indicator captures the level of JI in countries’ highest courts, which is used as a proxy for the independence of judges in the whole court system. Their data has recently been updated by Holsinger et al. (2017).

A completely different approach has been pursued by the World Justice Project (WJP). Whereas Linzer and Staton (2015) rely only on secondary data, the WJP relies exclusively on primary data. The primary data on which the index is built are assessments from local residents (1,000 respondents per country) as well as local experts. The current edition covers 113 countries to measure how the rule of law as experienced and perceived by the general public worldwide. Now, the rule of law is much broader than JI and JA, but some of their variables focus on both JI and JA. To make my point, I am including a list of the most relevant variables.

Box 1: JI-relevant variables as documented by the World Justice Project

- The government always obeys the decisions of the high courts, even when they disagree with these decisions
- In practice, the national courts in [contact("country")]) are free of political influence in their application of power
- The following question aims at identifying the main problems faced by the criminal courts in [contact("country")]. On a scale from 1 to 10 (with 10 being a very serious problem, and 1 being not a serious problem), please tell us how significant are the following problems faced by the criminal courts in the city where you live: a. Lack of independence of the judiciary from the government's power
- Based on your experience, out of all the cases in which the government had an interest (as a litigant or third party), in what percentage (%) of them did the government exercise undue influence to affect the outcome of the case?
- In practice, the national courts in [contact("country")]] are free of political influence in their application of power
- In practice, the local courts in [contact("country")]] are free of political influence in their application of power
- Assume that a government officer makes a decision that is clearly illegal and unfair, and people complain against this decision before the judges. a. Q11A. In practice, how likely is that the judges are able to stop the illegal decision?

Of the countries covered by the ENCJ paper, only Latvia, Lithuania, and Slovakia are not covered by the WJP. On the other hand, the WJP report includes a number of countries not included in the ENCJ paper, namely Albania, Bosnia & Herzegovina, the Czech Republic, Estonia, Georgia, Greece, Moldova, Serbia, Turkey, and Ukraine.

As a last example for readily available indicators of high quality, I want to mention a number of variables contained in the Varieties of Democracy (V-Dem) project. The goal of the project is to capture the many different dimensions of democracy over time. As of today, the dataset covers 201 countries with annual data from 1789 to 2017. V-Dem draws on advanced techniques for aggregating their 450 indicators. One chapter explicitly deals with the judiciary and contains aspects such as the number of judicial purges, whether there were any government attacks on the integrity of the judiciary, whether any court packing was observed, and whether judges are held accountable for their conduct. This dataset is fascinating not only because so many leading social scientists have contributed to it, but also because it has been produced to cover so many years.

Given these momentous and readily available indicators, the ENCJ attempt appears almost dwarfish. This is why the ENCJ should turn elsewhere.

Where to Add Value

Above, I already mentioned that the track record of judicial councils regarding the rule of law as well as the effectiveness of the judiciary seems to be rather problematic. A recent paper by Castillo Ortiz (2017), which is based on a survey carried out among judges by the ENCJ, adds even more doubts. It finds that in some European countries judges have a decidedly negative view on judicial councils. In Spain, for example, more than one third of all respondents claim that their judicial council has not been

---

3 Botero and Ponce (2011) describe the methodological approach of the WJP in more detail.
4 In a similar fashion, the WJP data allow the derivation of an indicator of JA. A more precise description can be found in Gutmann and Voigt (2018b).
respecting their independence. This is definitely neither the only nor the first paper finding that judicial councils can be a bane for judicial independence. It deserves particular mention here because it is based on data collected by the ENCJ.5

These findings should be alarming for the ENCJ. Although they should be alarming, they are not the last word on judicial councils. Most of the available studies rely on a rather small number of jurisdictions covered. The term judicial council is very broad and many different types are covered by the generic term. In their early contribution, Garoupa and Ginsburg (2009) propose to look at their composition (how many judges, how many politicians, how many lawyers etc.), the way in which they are appointed, their size, and their competences asking whether they have only administrative functions or also competences regarding the appointment, transfer, and disciplinary sanctions of judges. Bobek and Kosar (2014) propose to distinguish three emanations in Europe alone, namely a judicial council model, a courts service model, and a hybrid one.

It is not unlikely that particular traits of the organizational structure of judicial councils are responsible for most of their disadvantages. But this is merely a conjecture. It would be very important to know whether such a conjecture can be confirmed empirically. Such a project would need to comprise more judicial councils than those represented in the ENCJ. So I suggest that the ENCJ cooperate with other councils to inquire into organizational traits conducive to JI and JA.

Over the last decade or so, the influence of prosecutors has grown immensely in many countries. With regard to criminal cases, it is them who decide whether a case ever ends up in court. They are, hence, agenda-setters for judges, and they have been referred to as the ‘judge before the judge’ and ‘judge by another name’ (Weigend 2012, 378, 389). Prosecutors formulate charges, conduct the prosecution and argue the case in court. In order to take the strain off overloaded court systems, many jurisdictions have shifted powers and responsibilities from judges to prosecutors. As a result, some scholars consider prosecutors to be potentially the most powerful actors in the criminal justice system (Jehle and Wade 2006; Weigend 2012, 383–9). In a report entitled “The Disappearing Trial”, the NGO Fair Trials International (2017) offers information on the degree to which trial waver systems – a generic term for what is known as plea bargaining in the U.S. – are used today. Here is the percentage of cases for a number of European countries: Bosnia & Herzegovina 41%, Estonia 64%, Georgia 88%, Poland 43%, and Spain 46%.

In an early study, van Aaken et al. (2010) showed that the de facto independence of prosecutors is strongly correlated with corruption levels: the more independent the prosecutors, the lower the perceived level of corruption. It has even been argued that for politicians who are interested in reaching political goals via the judiciary, putting pressure on prosecutors – and, thus, tinkering with their independence – could be more attractive than trying to influence judges, simply because the latter would arouse a lot more opposition (Voigt and Wulf 2017). In sum, then, the de facto independence of prosecutors seems highly relevant for realizing the rule of law. Given its increasing relevance, it is amazing how little we know about it. There are many indicators on offer depicting JI, some regarding JA and prosecutorial independence, but none dealing with prosecutorial accountability. This is an important gap and it would be wonderful if the ENCJ decided to take on this important job possibly in cooperation with the International Association of Prosecutors or some other suitable organization.

5 Conclusion

Both judicial independence and judicial accountability are important. It is good that they receive considerable scholarly attention and have done so over the last 15 years. Several indicators purporting to measure them and to make them comparable across countries have been made available. Due to the wealth of already existing indicators, there is little value to adding yet another one. This comment argues that ENCJ supported research should focus on more pressing issues. Three such issues were named. They are (1) carrying out surveys among court users, (2) measuring organizational traits of judicial councils, and (3) developing an indicator depicting prosecutorial accountability.

References


5 Other papers critical regarding the effects of judicial councils include Haley 2006, Beers 2012, Popova 2012, Bobek and Kosar 2014.


COURTS & JUSTICE SOFTWARE THAT EXPANDS ACCESS TO JUSTICE

MODRIA IS THE WORLD’S LEADING ODR PLATFORM

Modria online dispute resolution expands access to justice by empowering people to resolve disputes 50 percent faster. It was created by the pioneers of ODR systems at eBay and PayPal and is the only solution that leverages Tyler’s proven ability to transform courts. Tyler serves more than 15,000 local government offices across all 50 states, Canada, the Caribbean, the U.K., and other international locations.

Learn more at tylertech.com.
Researching judicial ethical codes, or: how to eat a mille-feuille?

Elaine Mak

Abstract
Judicial ethical codes have become increasingly important in member states of the European Union as a point of reference for judges and as a means of securing public trust in judiciaries. In the ENCJ’s survey on the independence, accountability, and quality of judicial systems in Europe, sub-indicators were included to measure the existence of judicial ethical codes and relating training and supervision. This survey and the critical scholarly discussion about it provide valuable lessons for further empirical research on the perceptions of judges in EU member states relating to professional-ethical codes and standards. This article explores the most important lessons for this envisaged research. The main methodological considerations which come to the fore are: 1) to clearly delineate the research aims and central terms which are used in the survey; 2) to avoid insularity by including a representative and diverse group of respondents; 3) to validate and contextualise the results through a multi-method approach and critical debates with scholars and stakeholders.

Keywords: judicial ethics, judicial culture, empirical research design, ENCJ, European Union

1. Introduction
Ethical (or deontological) codes for legal professionals have been compared to a mille-feuille, the French pâtisserie composed of layers of puff pastry and pastry cream and topped with powdered sugar or icing. Indeed, the metaphor seems to work when considering the formal ‘crumbling’ and substantive heterogeneity which the set of available codes in the world displays. Crumbling, firstly, can be seen in the variety of codes on good conduct of legal professionals, the actors involved in developing and establishing these codes, and the different categories of legal professionals to which these codes apply. Written ethical codes for inter alia judges and lawyers have proliferated in the 21st century, both in national systems and internationally. These codes are developed within the professions, but increasingly also with the involvement of governments and legislators. Heterogeneity, secondly, concerns the variety in the content of ethical codes for legal professionals. Differences exist with regard to the specificity of codes, their guiding or binding nature (sometimes specified in the used terminology, e.g. code, guide, guiding principles), and the clarity of provisions. In this landscape of ethical codes, the guiding force for legal professionals and the internal coherence of norms are hampered by overlaps between codes as well as the inconsistent distribution of norms over internal guidelines and more formal types of regulation.

Despite the proliferation of ethical codes for judges in Europe, not much research has been conducted on the content of these codes and their influence on judicial functioning. Yet, professional ethics constitute one of the core elements of judicial culture, meaning “the features that shape the way in which the work of a judge is performed and valued within particular legal systems”. In this regard, judicial ethical codes are an important factor for achieving goals set by the European Commission for realising effective judicial cooperation in the European Union (EU) in the framework of the Lisbon Treaty. Against this background, this article aims to present a reflection on the methodological considerations for an envisaged empirical research on the perceptions

*Elaine Mak is Professor of Jurisprudence at Utrecht University and a researcher at the Montaigne Centre for Rule of Law and Administration of Justice. This article connects with a research project on European judicial culture(s), funded with a Vidi grant from the Netherlands Organisation for Scientific Research (NWO). Thanks to Erin Jackson, Niels Graaf, and Frans van Dijk for commenting on a draft version of this article and to the anonymous reviewers for their useful feedback. Any mistakes remain my own E-mail: e.mak@uu.nl

3 Moret-Bailly & Truchet, supra note 1, p. 10.
4 Ibid., pp. 32-44.
of judges concerning the influence of judicial ethical codes in the EU. In this reflection, lessons will be drawn from the design and results of the survey on judicial independence and accountability, which was conducted by the European Network of Councils for the Judiciary (ENCJ) in 2017. This survey and critical analyses from scholars will be consulted to answer the central research question for this article: which methodological considerations should be kept in mind when elaborating an empirical research design for studying the influence of judicial ethical codes in EU member states?

First, the topic of judicial ethical codes and its relevance in the EU context are further introduced (2). Following this, a brief overview of judicial ethical codes in selected EU member states will highlight main points of attention for a comparative and empirical research on these codes (3). Next, relevant parts of the EN CJ survey and a draft design for the envisaged new study on judicial cultures will be described as two examples of research design on judicial ethical codes, of which the combined results will be able to clarify the connections between formal arrangements for national judiciaries and the realisation of de facto independence and accountability (4). Finally, an analysis will be made of the relevant methodological considerations for researching judicial ethical codes, taking into account lessons which can be drawn from the EN CJ survey and the critical discussion of this survey (5). Some concluding remarks finalise the analysis (6).

2. Reasons for researching judicial ethical codes in the EU

Those who like the mille-feuille will know that it can be quite hard to eat this pastry gracefully. Likewise, finding one’s way in the multitude of ethical codes for the legal professions and dealing with substantive issues of clarity and coherence of norms can be a challenge for legal professionals themselves as well as for members of the wider society. Yet, tackling this challenge is of high importance for addressing changed societal demands. Transparency and accountability have become central values for ensuring public trust in professional performance, also for judiciaries, influenced by theories of new public management. Besides this, independence remains a prominent value for professionals and it is currently understood first of all as the professional’s neutrality with regard to (political) power.

These observations ring true also for the specific category of legal professionals formed by judges. Judiciaries are grappling with societal and organisational demands of effective, efficient, and client-oriented dispute settlement. At the same time, the independent position of judges vis-à-vis political powers requires constant maintenance in systems which strive to realise the ideal of the rule of law, understood here as the prevention of the arbitrary use of power. Against this background and with the aim of securing public trust, ethical codes for judges have been developed in many countries, in particular during the past two decades. Indeed, ethical codes can be a ‘sign board’ for society, which confirms that the core values of the society are taken seriously by judiciaries and individual judges, and these codes can express a set of concrete norms which may serve as a basis for accountability.

A further role of ethical codes emerges in the context of the EU, where national judges have responsibilities regarding the uniform application of EU law and the guarantee of equal legal protection for citizens. The responsibilities of judges have increased with the goals of judicial cooperation included in the Lisbon Treaty. In this regard, EU commissioner Věra Jourová has stated in her Foreword to the 2018 EU Justice Scoreboard:

Improving the effectiveness of justice systems is crucial for the respect of the rule of law. Without independent, high-quality and efficient justice systems there is no rule of law, no effective application of EU law, no business friendly environment and no mutual trust.
Yet, developments such as the rule-of-law crises in Poland and Hungary\(^{19}\) underline the societal and political challenges for realising cooperation based on mutual trust and recognition, for example with regard to the European arrest warrant in criminal cases.\(^{20}\) In this context, ethical codes have the potential of providing a normative common ground and concrete guidance for judges in EU member states to decide on professional-ethical dilemmas relating to the values of effective and efficient judicial protection and the realisation of the rule of law. In this regard, comparative influences from other EU member states and from European and international codes could inspire the content and strengthen the persuasive force of national ethical codes in specific EU member states and in this way assist judges in claiming professional neutrality vis-à-vis political actors in their country.\(^{21}\)

For judicial ethical codes to fulfil their functions properly, in particular to ensure access to justice and public trust in judiciaries, a sound formal structure and internal coherence of norms are required.\(^{22}\) Still, knowledge on the substantive similarities and differences between national judicial ethical codes in the EU and on the influence of these codes on judicial functioning is scarce.\(^{23}\) Pertinent questions for research then are: how have core values for judicial functioning, such as independence and accountability, been presented in judicial ethical codes in EU member states? How do judges perceive the influence of these codes for guaranteeing the core values of judicial functioning?

3. Judicial ethical codes in the EU: a mille-feuille indeed

A brief overview serves to demonstrate the diversity in judicial ethical codes in the EU with regard to the included core values and intended function of the code.\(^{24}\) In the overview in this section, a selection is presented of codes from four EU member states: the United Kingdom (England and Wales), France, Denmark, and Romania. These member states represent different legal traditions (common law and civil law),\(^{25}\) different professional cultures (consensus and honour),\(^{26}\) and different degrees of realisation of the ideal of the rule of law (established and ‘new’ liberal democracies).\(^{27}\) Based on this overview (3.1), some main characteristics of judicial ethical codes in Europe, which can serve as a starting-point for the design of empirical research, come to the fore (3.2).

3.1. A tour d’horizon

In England and Wales, a Guide to Judicial Conduct was first published in 2004 and updated several times since.\(^{28}\) This guide was established by a working group of the Judges’ Council and it has been authorised by the Lord Chief Justice of England and Wales and the Senior President of Tribunals.\(^{29}\) The guide applies to judges, coroners, and magistrates and it mentions three basic principles: independence, impartiality, and integrity.\(^{30}\) In the introduction to these principles, a distinction is made between the terms ‘guide’ and ‘code’ in order to emphasise the non-binding nature of the listed ethical norms. The purpose and status of the Guide are explained as follows:

> It is based on the principle that responsibility for deciding whether or not a particular activity or course of conduct is appropriate rests with each individual judge.

> This Guide is therefore not a code, nor does it contain rules other than where stated. Instead, it contains a set of core principles which will help judges reach their own decisions.

> In cases of difficulty or uncertainty, however, judges should always seek advice from the relevant leadership judge.\(^{31}\)

In France, the Conseil Supérieur de la Magistrature (CSM) established a deontological code for magistrates — encompassing judges as well as public prosecutors — in 2010.\(^{32}\) The CSM had a legal obligation to develop this *Recueil des obligations déontologiques*...
On the basis of the presented overview of judicial ethical codes in selected EU member states, some main characteristics of these codes can be identified. This overview also clarifies that the characteristics of these codes can be considered as sharing family

35 Conseil supérieur de la prud’homme, Recueil de déontologie des conseillers prud’hommes, mentioned by the Cour de cassation: <https://www.courdecassation.fr/venements_23/relations_juridictions_ordre_judiciaire_7108/commissions_nationales_discipline_8057/discipline_conseillers_8637/discipline_conseillers_38840.html> accessed 16 August 2018. See also art. L. 1421-2 and art. R. 1431-3-1 Code du travail.
36 CSM, supra note 32, p. X.
37 Ibid., p. X-XI.
38 Ibid., p. XII.
39 Ibid., p. XIII.
40 Ibid.
41 Ibid., p. XIV.
43 Ibid.
44 Council of Europe, Recommendation R(2010) 12 of 17 November 2010 on judges’ independence, effectiveness and responsibility, art. 73.
45 Association of Danish Judges, supra note 42.
47 Ibid., art. 2.
resemblances rather than one common conceptual template.\textsuperscript{48} The identified characteristics should be kept in mind in the design of empirical research on judicial ethical codes, as they clarify which stakeholders exist with regard to the codes and they highlight specific issues of content and effect of the codes. Moreover, this overview already highlights the importance of contextualisation of the outcomes of empirical research. The meaning and influence of judicial ethical codes in individual EU member states can only fully be understood when taking into account aspects of legal tradition, professional culture, and degree of realisation of the rule-of-law principle.

When considering judicial ethical codes as a specific category of norms, the following characteristics of these codes appear relevant for an empirical research: audience, content, type of norms, functions, and legal status.

**Audience: judges and others**

Ethical codes for judges apply to the active members of the profession. Some codes explicitly include retired judges.\textsuperscript{49} Sometimes, related professions are addressed as well. A striking aspect of the French code is that it does not apply to lay judges, for whom separate codes were developed. In this regard, the French code’s applicability to judges and public prosecutors relates to a specific elaboration of the rule-of-law principle in the constitutional status of magistrates and connected formal arrangements.\textsuperscript{50} The applicability of the English guide of judicial conduct to coroners can be explained by the coroner’s status as an independent judicial office holder, which has developed in the common law tradition.\textsuperscript{51}

**Content: universal values?**

The values expressed in the judicial ethical codes in the overview encompass the broad categories of independence, impartiality, professionalism, and integrity. The codes concern ethical conduct in the performance of professional tasks, such as conduct during court hearings or the deliberations on judgments. The codes also encompass broader conduct outside of the courtroom, for example political activity. Most of these codes have drawn inspiration from European and international documents, in particular the Bangalore Principles of Judicial Conduct developed under the auspices of the United Nations.\textsuperscript{52} Still, national contextual aspects transpire, for example in the French code’s emphasis on the dignity of the office of magistrate, which can be considered to reflect the professional culture of honour.\textsuperscript{53}

**Type of norms: values and standards**

Professional values present a frame of reference for proper judicial conduct. These values provide the normative underpinning for professional-ethical choices of judges.\textsuperscript{54} In a different category, professional standards for judicial performance are minimum norms for the adequate handling of court cases. These standards are connected with the idea of ‘judgecraft’, which encompasses the minutiae of procedural arrangements in courts.\textsuperscript{55} The content of these standards can range from court etiquette to break times or the processing of requests for adjournment of hearings and is intricately linked to experiences of procedural justice.\textsuperscript{56} In comparison, the codification of standards has developed less than that of values. The distinction between values and standards is important for understanding the functions and effects of judicial ethical codes, as normative points of reference (values) have a different influence than more practical instructions (standards). Still, a caveat is in order. Although values and standards can be distinguished conceptually, overlaps could occur in the practical elaboration of norms for judicial conduct in codes and other guiding documents. For example, both values and standards can mention continued education as a means of maintaining expertise and professionalism for the judgment of cases. A risk of this kind of overlap is that professional standards are included in judicial ethical codes as objective ‘measurable’ norms, which enable sanctions against judges who fail to meet these standards rather than help protect their independence.

**Functions**

Judicial ethical codes can have different functions. Most codes mention the intended functions and distinguish between the guiding role for judges (internal function) as well as the explanatory role for society (external function). A further nuance is possible when we consider the dynamics in the internal organisation of judiciaries and the position of judges in this organisational context. These

---


\textsuperscript{49} Judges’ Council, supra note 28, p. 5.

\textsuperscript{50} Art. 64 and 65 Constitution of 1958.


\textsuperscript{53} Cadet et al., supra note 26, p. 2.

\textsuperscript{54} Allard et al., supra note 14, p. 12.


\textsuperscript{56} Ibid., p. 319.
dynamics should be taken into account when researching the influence of ethical codes from the point of view of individual judges. The possible functions of codes can then be conceptualised as follows:\(^{57}\):

- Motivating function: incentive for judges to perform better in the judicial office based on concrete norms for good conduct;
- Accounting function: input for transparent and relevant disciplinary procedures, ‘sign board’ for society;
- Educational function: input for the education and training of new judges (mostly inductively, e.g. through dilemma training);
- Innovating function: space for adaptation of the judicial role to meet changing societal demands, based on a reflected solution of dilemmas.

As observed above, the accounting function becomes problematic if ‘measurable’ standards are included in judicial ethical codes. Then, the guiding nature of ethical norms as a means of strengthening judicial independence might be undermined by the use of these norms to sanction or fire judges.

**Legal status: guidelines or binding norms**

Judicial ethical codes can be presented as non-binding guidelines or as norms to which judges can be held accountable in internal evaluations or disciplinary proceedings.\(^{58}\) In the overview, the status of non-binding general guidelines prevails in states where the rule-of-law principle is firmly embedded in the professional culture. In states where the rule-of-law principle is not yet as firmly embedded, for example in post-communist states in Central and Eastern Europe, judicial ethical codes more often appear to have a stronger binding nature and contain more specific norms.\(^{59}\) This is visible in the example of the Romanian code, which encompasses specific and precise norms and explicitly mentions the role of the code in evaluations of judges and public prosecutors.\(^{60}\) In empirical research, this distinction should be kept in mind, as it can be expected to colour the perceptions of judges with regard to the influence of ethical codes on their professional conduct inside the courtroom as well as in their interaction with political actors and society.

4. Research designs: two examples

In this section, two approaches to studying judicial ethical codes will be described, followed by an analysis of the relation of these research designs to the questions and aims of the respective projects. The first example concerns the ENCJ survey on judicial independence and accountability, in which indicators on judicial ethical codes were included to measure the objective accountability of judges and judiciaries in 23 member and observer countries to the ENCJ (4.1).\(^{61}\) The second example concerns an empirical research project under development on the influence of judicial ethical codes in EU member states, which has the aim of assessing the possible alignment of professional-ethical values and practices of judges as a part of a European judicial culture (4.2).\(^{62}\)

4.1. ENCJ survey

The ENCJ’s survey has several aims.\(^{63}\) In the short run, the aim is to measure the de iure independence and accountability and perceived independence in countries in order to provide suggestions for the improvement of formal arrangements for the judiciary and judges. In the long run, the ENCJ aims to use the broad insights from the survey on the realisation of judicial independence in Europe to influence debates in countries where this principle is threatened, e.g. in Poland. Finally, the survey has the goal of giving a voice to judges, who can express their perceptions on the realisation of de facto judicial independence in their country.

The ENCJ’s approach consists in the development of indicators and sub-indicators for judicial independence and accountability on the basis of: 1) an analysis of available international principles and experiences of judges in the ENCJ with regard to independence and accountability; and 2) the listing of actual arrangements to aspire to in relation to these principles.\(^{64}\) The measurement through the survey is used to generalise outcomes in order to provide common aspiration levels and potential remedies for countries which do not yet meet these levels.\(^{65}\)

The ENCJ survey assesses formal independence and formal accountability of judges and judiciaries as well as perceived independence. The assessment of perceived accountability has not been included in the survey, presumably for practical reasons.\(^{66}\) In the survey, judicial ethical codes are addressed as an indicator of the objective accountability of the individual judge.\(^{67}\) This focus

---

58 Di Federico, supra note 2, p. 99; Simonis, supra note 23, p. 116-117.
59 But see Di Federico, supra note 2, p. 97, who relates this difference to the distinction between common law and civil law systems.
60 See above, par. 3.1.
61 ENCJ, supra note 8.
62 See supra note 7.
63 Van Dijk & Vos in this special issue, par. I. See also their lead article for further background information on the ENCJ’s survey.
64 Ibid., par. II.
65 Ibid.
66 Ibid., par. IV.1.
67 ENCJ, supra note 8, p. 17.
on accountability suggests that a connection is made to the function of ethical codes as a source of concrete norms and as a ‘sign board’ for society.68

Objective indicators in this survey concerned “the legal and other objectively observable aspects of the legal system that are essential for independence and accountability”.69 The sub-indicators measured for judicial ethical codes were: the existence of a code of judicial ethics; if existent, the public availability of this code; the availability of training on judicial ethics; and the existence or not of a responsible body to provide judges with guidance or advice on judicial ethics.70 For each of these sub-indicators, the developed questionnaire contained an option to answer with “yes” or “no”.71 The four sub-indicators on judicial ethical codes could together give a maximum score of 100%, with the result on each sub-indicator contributing 0% or 25% to the total score. The survey results show that scores on the four sub-indicators for judicial ethical codes yield a high average score of 85%. There is, however, a high variation between countries, with the lowest score being 0% and the highest score being 100%.72

For this measurement of objective indicators, a standardised questionnaire was filled out by the judicial councils or, in countries where no judicial council exists, by other governance bodies. The information obtained through this self-evaluation could have been retrieved also in other ways, that is: it “can be checked by anybody who is knowledgeable about the legal systems concerned”.73 According to the ENCJ, the objective indicators “set a standard about how formal arrangements should look like. ... For all indicators a high score is good and a low score bad.”74

As a general caveat, the authors of the ENCJ survey report warn that “the indicators have not been developed to create rankings of judicial systems, but can be used to discuss the strengths and weaknesses of judicial systems. Readers of the report are advised to treat the comparison of data from different countries with various geographical, economic and legal backgrounds with great caution.”75

4.2. Judicial cultures project

The judicial cultures project concerns an independent academic study on judicial cooperation in the EU, involving ethical, legal, and institutional aspects of judicial functioning.76 The final aim of this research is to establish whether a shared normative framework for the functioning of national judiciaries in the EU – in other words: a “European judicial culture” – can develop, which will enhance effective judicial protection and ultimately will contribute to the progressive realisation of the ideal of the rule of law in the EU. In the framework of this study, a survey is used to collect information on the perceptions which judges have about the functions and effects of judicial ethical codes in EU member states. These perceptions can include knowledge of and experience with codes as well as individual views on the role and usefulness of these codes.77 A distinction is made between professional values for judicial conduct and professional standards for judicial performance.78 This will enable assessing whether relevant differences between these categories can be identified, e.g. concerning the role of values and standards in external communication or in internal evaluation of judges, and the influence of codification (or not) of judicial ethical norms.

In this research, two hypotheses will be tested, which seem to reach further than the ENCJ’s inquiry into judicial ethical codes as an indicator of objective accountability.79 The first hypothesis is that judicial ethical codes and published professional standards can acquire a binding status through their use, e.g. through application in disciplinary procedures against judges or in internal individual evaluations of judges. The second hypothesis is that judicial ethical codes and published professional standards can be instruments which influence perceptions of de facto independence and accountability among judges or de facto independence and accountability as such (through a norm-setting effect).

The approach will be to develop indicators for professional-ethical performance of judges on the basis of: 1) a comparative analysis of the existence and content of judicial ethical codes and professional standards in selected EU member states; and 2) the listing of core elements of professional-ethical performance to aspire to on the basis of this comparative overview. The survey, which will receive a follow-up in interviews, aims to clarify how judges perceive these core elements and how they implement them in their daily practice. This analysis will enable drawing conclusions about the common ground (or not) between EU member states regarding the content of professional-ethical norms for judges and the ways of implementing these norms. Furthermore, this analysis will make it possible to assess the de facto effects of professional-ethical values and standards in realising judicial independence and accountability.

68 See above, par. 3.2.
69 ENCJ, supra note 8, p. 12.
70 Ibid., p. 17.
71 Ibid., p. 103 (questions 6a-6d).
72 Ibid., p. 20.
73 Ibid., p. 12.
74 Ibid., p. 19.
75 Ibid. See also Van Dijk & Vos, supra note 63, par. II.
76 See Mak, Graaf & Jackson, supra note 5.
78 See supra, par. 3.2.
79 See supra, par. 4.1.
The research design for this study (see Box 1) presents a bottom-up approach to the development of indicators, which are drawn from the developed values and practices in national systems. This is different from the ENCJ’s approach, which draws primarily from international principles.80 Moreover, this research does not have the aim, as the ENCJ’s project has, to improve judicial systems, criticise member states, or give voice to judges81. Its central aim is to investigate whether alignment of professional-ethical values and standards in the EU, as a basis for effective judicial cooperation, can develop.82 Of course, policy-makers, judicial networks and other stakeholders could use the results of this academic research to push their own agendas forward. Therefore, it will be important in the presentation of the results to emphasise the descriptive nature of the research and to distinguish its aim from the policy goals which stakeholders wish to address.

Box 1: draft design for the survey

Part I: Judicial ethical codes
Awareness of national code(s)
Consultation of national code(s)
Existence of professional debate on the applicable code(s)
Perception of disciplinary and evaluative effects of code(s)

Part II: Professional standards
Awareness of national standards
Consultation of national standards
Existence of professional debate on the applicable standards
Perception of disciplinary and evaluative effects of standards

Part III: European guidelines
Awareness of instruments developed by the EU, Council of Europe and Organisation for Security and Cooperation in Europe (OSCE) regarding judicial functioning
Opinion on instruments developed by the EU, Council of Europe and OSCE as a useful tool for realising professional-ethical values and standards
Perception of comparability and substantive alignment between codes and standards in EU member states

Part IV: General questions
Personal characteristics of respondents
Experience abroad and language proficiency

As a follow-up to the survey and interviews with judges, project seminars will be used to present the results to a broader audience and to involve different stakeholders in the reflection on the results. Furthermore, the results of the empirical research in this project will be connected with a content analysis of judicial ethical codes in order to validate the results of the survey and interviews. Contextualisation will be realised through a more close-up analysis of EU member states, in which the elements of legal tradition, professional culture, and the degree of realisation of the rule-of-law principle will be integrated in the research.

Still, some methodological challenges for the empirical research design remain, e.g. the concern of ensuring representative and trustworthy responses to surveys and interviews and the concern of combining the measurement of indicators with richer contextual studies in an accessible format. In the next section, these concerns will be addressed.

5. Methodological considerations

Based on the analysis in the previous sections, some methodological considerations for empirical research on the influence of judicial ethical codes in EU member states can now be mapped. These considerations concern two phases. The preparations for research, firstly, encompass the following aspects: determining the object and aim of the research, developing a conceptual framework, using a multi-method approach (5.1). The conducting of the research, secondly, requires attention for the validity and the contextualisation of results of an empirical survey (5.2).

80 Most international principles were developed with input from national judges. Therefore, overlaps between national and international ethical codes can be expected to be found.
81 See supra, par 4.1.
82 See supra, par. 2.
5.1. Preparations: optimising the research design

The initial question, of course, is: what will be measured in a survey and to what end? In this regard, research on judicial systems is characterised by the tension between a policy-oriented and an academic outlook. From a policy-oriented point of view, this research has the capacity to assist in addressing problems in the formal arrangements for judicial systems and problems in judicial functioning. From an academic point of view, more fundamental questions deserve attention as well, for example questions aimed at explaining specific aspects of judicial systems or questions aimed at supporting a normative reflection on the desired role of the judiciary in society. This more fundamental research does not always contribute to immediate problem-solving, but it can generate a better understanding of the foundations of judicial systems and insight into possible innovations for the future.

For the design of an empirical research, this tension between policy-oriented and academic outlooks entails that a choice is required. Research aimed at problem-solving can make use of a ‘light’ theoretical framework, in which existing normative values for judicial functioning are described and ordered. At the same time, the limitations of this kind of research for rethinking the foundations of judicial systems should be acknowledged. By contrast, fundamental research aimed at understanding or at stimulating normative debate can dive deeper into the interpretation and contextualisation of norms for judicial functioning, while setting more modest goals in terms of its contribution to legislative and policy reforms.

Against this background, the two presented research designs have complementary objects and aims. The ENCJ survey has the benefit of yielding a large set of objective data, which could signal ‘red flags’ – meaning a situation which requires urgent attention – in the realisation of core values for judicial functioning. The judicial cultures project will provide a deeper understanding of specific ways in which core values for judicial functioning are implemented. For each of these two outlooks on the topic, awareness of the limitations of the research is important in order to set realistic goals.

Furthermore, the analysis in this article reveals that researchers should be aware of the difficulties of working with indicators. As noticed elsewhere, awareness is required with regard to underlying assumptions, e.g. economic assumptions in the EU context, and political interests connected to the research. With regard to judicial ethical codes, there could be a risk of scores being used by governments or court managers to put pressure on judges, or by EU institutions to put pressure on member states to push through specific reforms. Another risk is that member states could use scores to downplay deficiencies of their systems. In this regard, it is important to observe that neither of the two surveys discussed in this article aims to rank the examined judicial systems.

Also, a precise definition of central terms is essential for generating results which can form the basis of a comparison between national judicial systems. Interdisciplinary perspectives from inter alia the social sciences (sociology, psychology), law, economics, and cultural anthropology can be used to design a sound conceptual and theoretical framework. Interestingly, the two presented surveys display a conceptual difference. The ENCJ’s survey considers judicial ethical codes as an element of objective accountability, which relates to the formal arrangements for judiciaries and for judges. By contrast, the judicial cultures project focuses on the perceived influence of judicial ethical codes, which connects with subjective independence and accountability. The combination of the results from these two research projects might, once the second project has been finalised, provide insights into the connections between formal arrangements for judiciaries and the realisation of de facto independence and accountability.

A final preparatory step concerns the choice of research methods. Quantitative empirical research (e.g. surveys) provides numbers, while qualitative empirical research (e.g. interviews) gives insight into the mechanisms behind the numbers. An important practical challenge for empirical research based on surveys or interviews concerns the realisation of a sufficiently high response rate and trustworthiness of responses. Also, researchers should be aware of the risk of a selective non-response bias, which means that only respondents with an interest or experience relating to the topic will participate. However, completed research projects which involved judges demonstrate that these risks can be overcome, e.g. by using a sufficiently large sample of respondents and by checking responses against information from other available sources. Of course, the data from surveys and interviews should be collected, processed, and stored in accordance with applicable norms of academic research ethics.

A further concern on the ENCJ’s survey design, which aims to serve practical goals, is that subjective indicators are not “actionable”. By contrast, a survey based on objective indicators can be replicated and might therefore have more credibility, e.g. for policy advice. In order to remedy this, it could be useful to measure how many times judges have been dismissed rather than to measure...
the perceptions which judges have of de facto independence.\textsuperscript{93} Still, perceptions can serve as a proxy for identifying problems with de facto independence in the examined countries. In the analysis, it should then be emphasised that perceptions are not experience. Also, information from other sources can be consulted to put the results in perspective.\textsuperscript{94} It is useful to already consider these possibilities and limitations in the preparatory phase of the research in order to ensure that the intended goals of the research can be achieved.

5.2. Conducting research: optimising the quality of the results

In the conduct of empirical research based on a survey a first concern is the validity of the results. For the ENCJ survey, the questionnaire on objective indicators was filled out by a “mixed bag”\textsuperscript{95} of organisations, as judicial councils have different compositions, competences, and resources. Moreover, in some countries a judicial council does not exist and for these countries the questionnaire was filled out by other governing bodies, such as ministries of Justice.\textsuperscript{96} In order to strengthen the trustworthiness of the research, the ENCJ has validated the results through checks performed by an internal expert group.\textsuperscript{97}

Some further considerations can be kept in mind with an eye to improving the validity of survey research. The external validation workshop on the ENCJ survey\textsuperscript{98} yielded the following suggestions:

- Case studies can provide richer information on the situation in different countries. As a means of selection, the method of grouping of indicators (clustering) would make it possible to identify “red flags”\textsuperscript{99} in specific countries. An example of clustering with regard to judicial systems could concern indicators on corruption, the number of disciplinary proceedings against judges, and the length of proceedings. A case study could then focus on countries where these topics in combination seem problematic.
- A longitudinal analysis for each country can provide further insight into the development of the values of judicial independence and accountability.
- The combination of a survey with interviews enables checking the validity of the survey results as well as collecting more in-depth information for explaining the results.
- It is important to avoid insularity in the collection of data. This can be realised by asking the same questions to different stakeholders to bring out inconsistencies in answers. Categories of respondents which could be added when studying judicial systems include lawyers, court users, and politicians.

Finally, contextualisation is an important but challenging aspect of comparative research, which also requires attention in research on judicial systems. For the ENCJ survey, critical comments concerned issues relating to the use of indicators, such as the risk of contingent circumstances being moved back in indicators or cultural differences not being acknowledged in the comparative analysis of countries.\textsuperscript{100} Remedies can be to relate institutional design of countries to historical developments and to conduct a contextual analysis for each country, in which cultural and other aspects are taken on board.\textsuperscript{101}

\textsuperscript{93} Ibid.
\textsuperscript{94} See supra, par. 5.2.
\textsuperscript{95} Term used during the workshop in Utrecht, see supra note 9.
\textsuperscript{96} See supra, par. 4.1.
\textsuperscript{97} Van Dijk & Vos, supra note 63, par. IV.
\textsuperscript{98} See supra note 9.
\textsuperscript{99} See supra note 9.
\textsuperscript{100} Comments made during the workshop in Utrecht, see supra note 9.
\textsuperscript{101} See further M. Siems, Comparative Law, Cambridge: Cambridge University Press 2014.
The discussed methodological aspects of research on judicial ethical codes in the framework of the ENCJ survey and the judicial cultures project have been summarised in Table 1.

<table>
<thead>
<tr>
<th>EN CJ survey</th>
<th>Judicial cultures survey</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Object</strong></td>
<td></td>
</tr>
<tr>
<td>Existence of a (published) judicial ethical code, training and supervision</td>
<td>Awareness, use, and opinion on judicial ethical codes</td>
</tr>
<tr>
<td><strong>Objective accountability (formal arrangements)</strong></td>
<td><strong>Subjective independence (perceptions)</strong></td>
</tr>
<tr>
<td><strong>Aim</strong></td>
<td></td>
</tr>
<tr>
<td>Diagnosis and problem-solving</td>
<td>Understanding and normative reflection</td>
</tr>
<tr>
<td><strong>Concepts</strong></td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>Codes and standards</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td></td>
</tr>
<tr>
<td>Collect general knowledge</td>
<td>Measure perceptions Evaluation by judges</td>
</tr>
<tr>
<td>Self-evaluation by judicial councils</td>
<td>(Comparative) content analysis of codes Connect results with legal tradition, professional culture, degree of realisation of rule-of-law principle</td>
</tr>
<tr>
<td><strong>Validity</strong></td>
<td></td>
</tr>
<tr>
<td>Combine with check by academics</td>
<td>Combine with interviews</td>
</tr>
<tr>
<td><strong>Contextualisation</strong></td>
<td></td>
</tr>
<tr>
<td>Idea: develop case studies, e.g. on ‘red flags’ Idea: develop a longitudinal analysis for countries Connect results with cultural differences</td>
<td>(Comparative) content analysis of codes Connect results with legal tradition, professional culture, degree of realisation of rule-of-law principle</td>
</tr>
<tr>
<td><strong>Main challenge</strong></td>
<td></td>
</tr>
<tr>
<td>Trustworthiness of self-evaluation</td>
<td>Representativeness of sample of respondents</td>
</tr>
</tbody>
</table>

Table 1: methodological aspects of empirical research on judicial ethical codes

6. Conclusion

This article set out to investigate which methodological considerations should be kept in mind when elaborating an empirical research design for studying the influence of judicial ethical codes in EU member states. In the ENCJ’s survey on the independence, accountability, and quality of judicial systems in Europe, sub-indicators were included to measure the existence of judicial ethical codes and relating training and supervision. This survey and the critical scholarly discussion about it provide valuable lessons for further comparative and empirical research on the perceptions of judges relating to professional-ethical codes and standards. For the envisaged research in the judicial cultures project, which was presented in this article, the most important lessons can be summarised as follows: 1) to clearly delineate the research aims and central terms which are used in the survey; 2) to avoid insularity by including a representative and diverse group of respondents; 3) to validate and contextualise the results through a multi-method approach and critical debates with scholars and stakeholders. The results of this upcoming research will hopefully contribute to a better understanding of the mille-feuille of judicial ethical codes in the EU and a reflection on a recipe (but not a bake-off!) for the future.

Literature


J. Bell, Judiciaries within Europe: A Comparative Review, Cambridge: Cambridge University Press 2006


M. Fabri, Methodological Issues in the Comparative Analysis of the Number of Judges, Administrative Personnel, and Court Performance Collected by the Commission for the Efficiency of Justice of the Council of Europe (2017) Oñati Socio-Legal Series 7(4) pp. 616-639


J. Moret-Bailly & D. Truchet, Pour une autre déontologie des juristes, Paris: PUF 2014

G.Y. Ng, Quality of Judicial Organization and Checks and Balances, Antwerp: Intersentia 2007


M. Siems, Comparative Law, Cambridge: Cambridge University Press 2014


J. Vranken, Exciting Times for Legal Scholarship (2012) Law and Method (February)

Pitfalls in data gathering to assess judiciaries
Marco Fabri

Abstract
This paper is divided into two parts plus some concluding remarks. The first one deals with some problems in comparing the number of judges, court personnel, and caseflow in European judiciaries. Data come from the Commission for the Efficiency of Justice (CEPEJ) of the Council of Europe. The second part deals with some pitfalls in the data gathering carried out by the European Network of Councils for the Judiciary (ENCJ) in the attempt to measure judicial independence and accountability. Each case study brings some hints, summed up in the concluding remarks, that may be useful to improve both exercises.

Keywords: Comparative judicial systems, Judicial independence and accountability, Commission for the Efficiency of Justice (CEPEJ), European Network of Councils for the Judiciary (ENCJ).

1. Introduction
“When you can measure what you are speaking about, and express it in numbers, you will know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the state of Science, whatever the matter may be” (Kelvin 1883, p. 73).

Maybe this is a little extreme. This famous quote comes from a physicist who also stated that: “I have not the smallest molecule of faith in aerial navigation other than ballooning” 2, however the issue to measuring phenomena not only in the so-called hard sciences, but also in the social sciences is of paramount importance.

Court’s values and performances are not exceptions. They are difficult to measure, but in recent years some more attention has been paid to try to say something about them. This is due to several reasons, such as, just to mention a few: a more managerial approach following the “new” public management wave with more pressure for better courts’ performance and the assessment of judges’ productivity, the enlargement of the European Union and the hope to share common values and similar performance, the economic crisis, the introduction of performance-based-budgets, the excessive length of the judicial proceedings and the need to understand the reasons behind it.

Measuring cannot disregard reliable data, which also are the basis for an informed fact based policy making at the supranational, national, or court level.

This paper analyses two remarkable efforts to collect data on courts’ functioning and values. The first part of this work addresses the comparability of some data coming from the most comprehensive data collection on the functioning of European judiciaries carried out by the Commission for the Efficiency of Justice (CEPEJ) of the Council of Europe. 3 These data are often used to make comparisons across countries, and they are also used in the European Union Justice Scoreboard. The second part takes into consideration the work carried out by the European Network of Councils for the Judiciary (ENCJ) on judicial independence and accountability, which has been finalized in the 2017 Report: Independence, Accountability and Quality of the Judiciary. This report is also at the basis of the paper by Frans Van Dijk and Geoffrey Vos (2018), which is the backbone of this special issue. 4

1 Marco Fabri is Senior researcher at the Research Institute on Judicial Systems, National Research Council of Italy (IRSIG-CNR), Via Zamboni, 26, 40126 Bologna e-mail: marco.fabri@irsig.cnr.it. IRSIG’s mission is to study and carry out research in the domain of justice, both as a scientific endeavour and as a service for policies to improve the administration of justice. IRSIG’s scientific effort is also directed at facilitating the convergence of basic features of European Union justice systems. IRSIG’s research deals with the “law in action”, governance, court management and ICT, working practices, design and assessment of judicial institutions and reforms in comparative perspective. www.irsig.cnr.it.


3 The first part of this paper has also been presented at the seminar “Too few judges? Regulating the number of judges in society”, held in Onati the 30th of June 2016, which brought to the publication: *Methodological issues in the comparative analysis of the number of judges in European countries*, Onati Socio-Legal Series, v.7, n. 4.

4 Frans van Dijk and Geoffrey Vos, A Method for Assessment of the Independence and Accountability of the Judiciary, International Journal for Court Administration, Volume 9, nr. 3, 2018, p.1-21. The ENCJ’s report also deals with a section on “quality”, which will not be dealt with in this paper.
The CEPEJ and ENCJ exercises are different in many ways, but they have some similarities in the methodological approach that makes it interesting to analyze them both, in particular as far as the measurement problems are concerned. In the concluding remarks I will emphasize the importance of reliable datasets for making meaningful comparisons, and I will give some suggestions for improving both exercises. Sophisticated and fancy statistical or economic analysis can be fascinating, but they can be dramatically wrong due to the poor reliability of the data set that they may use. Comparative quantitative analyses in the judicial context should be approached carefully and humbly. They usually need to be supported by qualitative analyses that can improve the interpretation of data.

2. CEPEJ data on judges and court staff

The Committee of Ministers of the Council of Europe established the Commission for the Efficiency of Justice (CEPEJ, Commission européenne pour l’efficacité de la justice) in 2002 “for improving the quality and efficiency of the European judicial systems and strengthening the court users’ confidence in such systems.” CEPEJ’s mission is to propose pragmatic solutions as regards judicial organization, to enable a better implementation of the Council of Europe’s standards in the justice field, to contribute toward relieving the caseload of the European Court of Human Rights by providing States with effective solutions to prevent violations of the right to a fair trial within a reasonable time.

CEPEJ’s work is organized in plenary sessions and ‘working groups’ that deal with some specific issues. In particular, one group deals with the “Evaluation of European Judicial Systems”, which collects quantitative and qualitative data on several topics of the judicial systems of the Member States. The collection is carried out through a questionnaire with more than 200 questions, which is filled in by the Member States national correspondents every two years. This is a unique collection of data and information about the functioning of European judicial systems since 2004, which has no equal in any other study carried out by international organizations or researchers.

The data collection organized by CEPEJ also is the basis for the ‘European Union Justice Scoreboard’. The Scoreboard is published every year since 2013 and compiles data only from the European Union countries. The Scoreboard is “an information tool aiming at assisting the EU and the Member States to achieve more effective justice by providing objective, reliable and comparable data on quality, independence and efficiency justice systems in all Member States”. (European Commission 2016, p. 1). Data on judges, court personnel, resources, and performance are usually provided for by CEPEJ.

Over the years, the CEPEJ Evaluation working group, the secretariat, and the countries’ national correspondents have constantly worked together to improve the reliability and consistency of the data collected. A quite detailed ‘Explanatory note’ has been drafted, and periodically amended, to “assist the national correspondents and other persons entrusted with replying to the questions” and, in so doing, to ensure that concepts are addressed according to a common understanding.

Country replies to the questionnaire often come with further comments that are very informative and valuable to give a better interpretation of the numbers. This allows grasping several important features of the different systems. However, as I will discuss later, quite often, they are still not good enough to make ‘safe comparisons’ across countries.

Some other collections of data also have been carried out in the criminal field but, as of today, the most current data on the functioning of European judicial systems is the CEPEJ collection.

The CEPEJ reports over the years are full of warnings to avoid superficial comparisons or, even worse, meaningless ranking of the Council of Europe Member States. In CEPEJ’s words: “Comparing quantitative figures from different States or entities, with different geographical, economic, and judicial background is a difficult task which must be addressed cautiously […] Data cannot be read as they are but must be interpreted in the light of the methodological notes and comments. Comparing is not ranking” (CEPEJ 2016, p. 6).

---

5 This was the topic that it was asked me to address during a workshop to debate and assess the ENCJ 2017 Report. The workshop was organized by the ENCJ and the University of Utrecht; it was held in Utrecht, The Netherlands, 12-13 April 2018.
6 CEPEJ has 47 members appointed by the 47 Member States, and it is supported by a Secretariat within the Directorate General of Human Rights and Legal Affairs of the Council of Europe.
7 http://www.coe.int/t/dghl/cooperation/Cepej/presentation/CEPEJ_depliant_en.pdf
8 http://www.coe.int/t/dghl/cooperation/Cepej/presentation/CEPEJ_depliant_en.pdf
9 Another working group is ‘The Saturn Centre for Judicial Time Management’, which is charged with collecting data and information about the length of judicial proceedings in the Member States, sharing practices, and developing tools and innovative ideas to improve the pace of litigation, and to prevent violations of the right for a fair trial within a reasonable time.
10 See also http://ec.europa.eu/justice/newsroom/effective-justice/news/160411_en.htm
In this large data collection from different countries, consistency is a constant problem, and many efforts have to be done to define a clear and common ‘unit of analysis’ of what is required to be counted. Over the years, in the constant attempt to improve the consistency of the data collected, several definitions have been refined during the meetings among the national correspondents. Usually, these definitions are the result of a difficult process of including a large variety of differences in common categories concerning the legal system in the Member States.

Only an informed reading of the figures collected can avoid flaws; data cannot be passively taken but must be interpreted in the light of qualitative information that can better explain the meaning of that number. Large comments sometimes included in the replies are useful and informative for a better understanding of ‘cold figures’.

Among the data collected there are the number of judges and court personnel employee in each judiciary. CEPEJ defines, a ‘judge’ as “an entrusted person with giving, or taking part in, a judicial decision opposing parties who can be either natural or physical persons, during a trial” (CEPEJ 2016, p. 81).

Figures about judges are collected by dividing them into three categories: Professional judges “those who have been trained and who are paid as such”, and whose main function is to work as a judge and not as a prosecutor. The fact of working full-time or part-time has no consequence on their status; Professional judges practicing on an occasional basis “paid as such”, Non-professional judges “volunteers who are compensated for their expenses and who give binding decisions in courts” (CEPEJ 2016, p. 81).

Judges, and other court personnel are counted using the “Full-Time Equivalent” (FTE) method, to try to ensure consistent data and provide a starting point for any comparative analysis.

Unfortunately, the FTE method does not seem to have been applied by all the countries. As the CEPEJ report states: “only some states have indicated details (judges seconded to the ministries, judges on maternity leave, for instance)” (CEPEJ 2014, p. 156). Doubts about the appropriate use of the FTE method are evident from the comments about the data supplied by several countries. Inconsistency in the application of the FTE counting of judges can generate serious problems in any comparative analysis.

Another problem is related to the structure of the different European judiciaries and their jurisdictions. Courts’ jurisdiction can be quite different in the various countries, and this can affect the counting of both judges and other court personnel significantly. Usually, there are four major jurisdictions: civil, criminal, administrative, financial/tax. Generally speaking, in many countries civil and criminal matters are called ‘ordinary jurisdictions’, dealt with by ‘ordinary courts’ which can be internally divided into two or more specialized branches or sectors. Administrative and tax matters can be quite often two autonomous jurisdictions dealt with by different specialized courts, with dedicated judges and court personnel.

The jurisdiction considered in the counting of judges and court personnel are not always specified neither in the data collected nor in the comments. In some cases, countries have interpreted the counting in different ways. For example, France included in the total number of judges the administrative judges, while they have not been counted in Italy, although both countries have similar jurisdictions for administrative matters.

The three sub-categories of judges (“professional judge”, “professional judges sitting occasionally”, “non-professional judges”) also raise some concern, because they have not been interpreted consistently. For example, in the 2014 report, Germany also included as ‘professional judges’ the number of professional judges sitting in courts part-time occasionally, and it is not clear whether they have been calculated using the Full-Time Equivalent.

The further problem is that in some countries, ‘specialized’ civil or criminal cases (e.g. labor, small claims, commercial, misdemeanor) can be dealt with by different kinds of judges, and it is not always clear if and how these judges have been counted. For large countries, this can alter the numbers reported significantly and jeopardize any cross-country analysis, unless an in-depth qualitative analysis is carried out.

For example, in France, the numerous commercial and labor cases are not dealt with by the civil courts but by specialized courts with ‘judges/adjudicators’ appointed by the business community. In France, the number of ‘non-professional judges’ does not include ‘jurors or lay judges’; however in other countries such as, for example, Germany and Slovenia jurors or lay judges were included in the numbers. Norway reported both in 2012 and 2014, the same number of 43,000 “non-professional judges”, but it is not explained why they are so many. Several others examples can be given from other countries on the same note.

Also, non-professional judges are indicated in gross numbers and not in Full-Time Equivalent. It could happen that a non-professional judge works only a few hours per year, whereas others can serve almost full time.

---

13 “The full-time equivalent indicates the number of persons working the standard number of hours; the number of persons working part time is converted to full-time equivalent. For instance, when two people work half the standard number of hours, they count for one “full-time equivalent”, one half-time worker should count for 0.5 of a full-time equivalent” Cepex (2013), Explanatory Note to the Scheme for Evaluating Judicial Systems. 2014-2016 Cycle, Strasbourg, France, p. 2.

14 The CEPEJ Explanatory note to the scheme for evaluating judicial systems 2014-2016 cycle, CEPEJ (2015), p. 12, it is quite clear: “non-professional judges are those who sit in courts (as defined in question 46) and whose decisions are binding but who do not belong to the categories mentioned in questions 46 and 48 above. This category includes namely lay judges and the (French) ‘juges consulaire’. Neither the arbitrators, nor the persons who have been sitting in a jury (see question 50) are subject to this question”, but not all the National correspondents used this definition.
As shown, if it is difficult to count the number of judges in a consistent way across the different countries, the counting of “non-judges” is even more difficult.

The CEPEJ collects data about the ‘Rechtspfleger’\textsuperscript{15} in the countries that have such a position.\textsuperscript{16} Data are also collected on the number of ‘judicial advisors or registrars’, and ‘administrative staff’, however, these categories are not always very easy to differentiate from the other two categories listed in the CEPEJ questionnaire: ‘technical staff’, and ‘other kind’ of staff.\textsuperscript{17} It is also not clear in which category the ‘law assessors’ or ‘law clerks’ are reported to fit. They are judges’ assistants, assigned with for example legal research or drafting court decisions. Such functionaries are employed in several countries such as the Nordic countries and Switzerland.

To sum up, on the one hand, CEPEJ data collection on the number of judges and court personnel is an outstanding effort and it gives a first interesting overview of the different judicial settings and basic figures. On the other hand, some major problems in the comparability of data have to be mentioned. In particular: a) FTE counting does not seem to be used consistently by all the reporting countries; b) the figures on judges can mix up different jurisdictions which are not consistent across countries; c) the counting of judges in the three categories proposed (professional, professional practicing on occasional basis, non-professional judges) is not consistent across countries, d) all these problems are even more severe in the counting of “non-judges”.

3.CEPEJ data on caseflow

Courts’ caseflow data are collected for incoming, pending, and disposed cases. These data are the basis for three basic indicators of court functioning: Clearance rate (disposed cases/incoming cases x 100); Case turnover ratio (disposed cases/pending cases x 100), Forecast disposition time (365/case turnover ratio) (CEPEJ 2014, p. 191 and CEPEJ 2016, p. 185).

Also for this data collection, the main concern is to have a clear definition about incoming, disposed, and pending cases, and it is not that simple as it might appear. There is not a common consolidated definition of what a civil case or a criminal case is, as well as there is not a commonly shared interpretation of disposed and pending cases across countries.

Data on civil cases are also collected for seven case categories: civil and commercial litigious; civil and commercial non-litigious; non-litigious enforcement; non-litigious land registry; non-litigious business; administrative law cases; other cases.

Once again, the different settings of European judicialities generate some significant problems. For example, in Austria, court statistics do not allow a distinction between litigious and not-litigious cases, and then figures supplied are just an estimation. In the Netherlands, it is not possible to make a distinction between litigious and non-litigious incoming cases, but this is only possible after the case has been disposed of. The Czech Republic reported the number of electronic payment orders in the category ‘other cases’, while other countries in the ‘non-litigious civil and commercial’. In Norway, courts also have functions of public notaries and marriages that have been estimated at 25,000 per year. These cases have been included in the ‘other cases’ category. In Lithuania and Denmark, administrative cases have been included in the ‘other cases’ category. The same category has been used for insolvency registry and labor cases in Hungary, but not in several other countries.

In general, it is not always clear if and in which category administrative cases have been counted. In some countries, civil cases can include administrative cases, while in others they are counted in a different category.

\textsuperscript{15} As reported in the CEPEJ report 2014 (p. 175), the Rechtspfleger is “an independent judicial body, anchored in the constitution and performing the tasks assigned to it by law. [...] The Rechtspfleger does not assist the judge, but works alongside the latter and may carry out various legal tasks, for example in the areas of family and guardianship law, the law of succession, the law of land registry and commercial registers. He/she also has the competence to making judicial decisions independently on granting nationality, payment orders, execution of court decisions, auctions of immovable goods, criminal cases, the enforcement of judgments in criminal cases (including issuing arrest warrants), orders enforcing non-custodial sentences or community service orders, prosecution in district courts, decisions concerning legal aid, etc.; the Rechtspfleger, to a certain extent, falls between judges and non-judge staff, but does not have the status of judge.”

\textsuperscript{16} 16 countries reported to have some kind of staff close to the definition given to Rechtspfleger, these countries are mainly in central Europe and the Balkans, in somehow was inspired by the German tradition.

\textsuperscript{17} Explanatory note to the scheme for evaluating judicial systems 2014-2016 cycle, CEPEJ (2015)\textsuperscript{2}, p. 13 “Non-judge (judicial) staff directly assist a judge with judicial support (assistance during hearings, (judicial) preparation of a case, court recording, judicial assistance in the drafting of the decision of the judge, legal counselling - for example court registrars). If data has been given under the previous category (Rechtspfleger), please do not add this figure again under the present category. Administrative staff are not directly involved in the judicial assistance of a judge, but are responsible for administrative tasks (such as the registration of cases in a computer system, the supervision of the payment of court fees, administrative preparation of case files, archiving) and/or the management of the court (for example a head of the court secretary, head of the computer department of the court, financial director of a court, human resources manager, etc.). Technical staff are staff in charge of execution tasks or any technical and other maintenance related duties such as cleaning staff, security staff, staff working at the courts’ computer departments or electricians. Other non-judge staff include all non-judge staff that aren’t included under the categories 1-4”.

70 International Journal for Court Administration/December 2018
Comments reported by the countries in the CEPEJ reports show how hard it would be to make any comparison about the figures reported in each category based on the different definitions and then the counting, given by each country.\(^{18}\) These few examples show that it is misleading to make comparisons across countries without considering the specificity of the court’s performance figures. This can be even more misleading if courts’ performance is put in relation to personnel resources, due to the difficulties in comparing the number of judges and court staff mentioned in the previous section.

In addition, even though the reliability of data has improved, there are still some concerns about the quality of the data collected. Some checks in each country should be done to verify the correspondence to what is supposed to be reported and what is done in practice, also to assess the possible error rates.

Therefore, notwithstanding the remarkable efforts carried out by CEPEJ, there are still some significant problems in the collection of data about the number of judges, court personnel, and caseflow, which jeopardize a reliable comparative analysis across countries.

4. The ENCJ’s exercise to measure judicial independence and accountability

The second part of this paper deals with the ENCJ’s recent exercise to measure judicial independence and accountability in the EU Member States through a scorecard. In particular, I will focus, as requested by the editors of this special issue, on some pitfalls related to the measurement method used to assess the variables ‘independence’ and ‘accountability’.

The ENCJ “unites the national institutions in the member States of the European Union which are independent of the executive and legislature, and which are responsible for the support of the judiciaries in the independent delivery of justice”. The strategic objectives for 2018-2021 are to “provide support for the independence, accountability, and quality of judiciaries in Europe [...] to promote access to justice in a digital age [...] to strengthen mutual trust among the judiciaries in Europe”.\(^{19}\)

Among the 2016-2017 ENCJ activities, a specific group dealt with the challenge to measure judicial independence and accountability in each participating country with a scorecard. The scorecard is calculated by adding the points that each judiciary scored on several indicators and sub-indicators. The scorecards, according to the ENJC, should stimulate judiciaries and governments to policies’ design and implementation on judicial independence and accountability.

The 2017 report by the ENJC, along with the paper by Van Dijk and Vos, are the backbone of this special issue. Van Dijk and Vos’ paper is based on the 2017 ENCJ report; but it further explains the approach in a broader theoretical framework.

The work carried out is now in the planned phase to be externally reviewed by the scientific community, and by ENCJ’s partners (ENCJ 2017, p. 7). This paper is part of this external review.

The ENCJ work is mainly based on two questionnaires designed, submitted, and analyzed by the ENCJ. The first one was submitted to Judicial Councils, or other administration of justice bodies, and it was aimed to collecting data on the formal legal setting of judicial independence and accountability. The second questionnaire aimed to gather data on the judges’ self-perceptions of their independence. It was submitted to EU judges, through the national Judicial Councils. 11,712 judges from 26 countries participated in the survey (ENCJ 2017, p. 4 and 34).\(^{20}\)

Data from other European surveys were also used to calculate an indicator about the citizens’ perception of independence. Quite interestingly: “The correlation between this indicator and the perceived independence by judges is high, showing that the perceptions of judges of their actual independence are fairly in agreement with those of citizens” (ENCJ 2017, p. 21).

The ENCJ effort is different from the CEPEJ’s exercise for objectives, magnitude, and resources, but they have quite a few similarities. For example, it is similar the approach to data collection. CEPEJ employs national correspondents, who are mainly members of Ministries of Justice or dedicated Courts’ agencies (Northern European model). ENCJ has collected data through Judicial Councils, or other institutional bodies including sometimes Ministries of Justice in countries without a Council.

Both institutions share the same concern about the use of the data collected. ENCJ has produced “country profiles” with the data collected, and it warns that they “must be used with circumspection, due to the unavoidable arbitrariness of some categorizations and scoring [...] determining what is good and what is less good practice is based on shared values and ideas within the ENCJ, and such is not absolute science. Still, the profiles need to be taken seriously to set priorities for change” (ENCJ 2017, p. 84).

---

\(^{18}\) Cepej (2014), p. 219-220 “Austria: misdemeanours and/or minor criminal cases include all offences fined or punish with a prison sentence up to one year and must not be decided by a jury [...] Belgium: severe criminal offences include cases dealt with by first instance ordinary criminal courts. Misdemeanours/ minor criminal cases include cases dealt with by the Police Court [...] Denmark: severe criminal cases are defined as those cases where a lay assessor participates or cases dealt with by a jury; non-contested plea cases (guilty plea) are included as severe criminal cases. Misdemeanours and/or minor criminal cases are typically cases where the maximum sentence is a fee [...] Poland: misdemeanours cases includes the offences punishable by a maximum penalty up to 1 month of detention or a fine (or both). All other criminal cases are severe cases. Statistics contain also the so called ‘organisation cases’ which do not deal directly with crimes”. See also comments reported in the 2016 dynamic data-base (2014 data).

\(^{19}\) ENCJ members are 23 national institutions. There are also 14 observers Institutions, mainly Ministries of justice and Court Administration organization of countries that do not have a judicial council. https://www.encj.eu/ last visited August 2018.

\(^{20}\) This exercise is planned to take place every two years. (ENCJ 2017, p. 7).
As it is for CEPEJ: “The indicators have not been developed to create rankings of judicial systems, but can be used to discuss the strengths and weaknesses of judicial systems. Readers of the report are advised to treat the comparison of data from different countries with various geographical, economic and legal backgrounds with great caution” (ENCJ 2017, p. 19).

The answers collected have been codified by the ENCJ working group with a score for each ‘option’ (sub-indicator), following some self-elaborated guidelines. The scorecard for each judiciary is the sum of the points of each indicator.21

More in detail, scores in the questionnaire answered by Judicial Councils have different minimum and maximum (e.g., 0-1, 0-3, 0-5, 0-10, 0-15), depending on the indicator. Therefore, “for all the indicators a high score is good and a low score is bad” (ENCJ 2017, p. 19). For example, question 8e “Can the management of the court exert pressure in individual cases on the way judges handle their cases with respect to the timeliness/efficiency of judicial decisions?” scores 3 points if the answer is ‘no’, it scores 0 points if it is ‘yes’. Question 7a “Can a judge be transferred (temporarily or permanently) to another judicial office (to other judicial duties, court or location) without his/her consent?” Scores 15 points if the answer is ‘no’, it scores 0 points if the answer is ‘yes’.

These examples raise two concerns about the questionnaire submitted. The first one is more specifically related to the formulation of the questions. The above examples, for instance, ask such fundamental questions as whether judges are free from pressure or not. In both cases, a self-compilation of the scorecard is likely to lead to an improper assessment of the judiciary. One concern is the legitimacy of these scores and the related final score of each judiciary. Another concern is the reliability of some answers.

As already mentioned, also this data collection relies on self-compilation, which raises some concerns about the genuine, and then reliability, of some answers.

Van Dijk and Vos are aware of this risk (p. 23): “In some situations, this may lead to self-serving bias. This is difficult to avoid, but it is necessary within the ENCJ framework, who may help to improve the quality of the questionnaire and to check the consistency of the answers.

A couple of specific examples from the Italian case, can be useful to show the pitfalls that may be due to the self-compilation of the questionnaire by Judicial Councils without any further check.

Some questions are just related to legal matters about the judiciary and, in this respect, the margins of interpretation are limited, although still possible. For example, question 7a “Can a judge be transferred (temporarily or permanently) to another judicial office (to other duties, court or location) without his/her consent?” was answered “yes” by the Italian Judicial Council. However, as far as I know, the cases in which this could happen are exceptional and very rarely occur (Contini et al. 2017, Di Federico 2012 and 2005). The answer given by the Council was just legally oriented, missing the ultimate issue raised by the question.

Other questions asked to make an assessment about different issues. In these cases, answers are more discretionary, and maybe questionable, as it is the related score of the indicator. For example, question number 3a asked: “Is the funding of the judiciary sufficient as to allow the courts to handle their caseload” (and other items). The Italian Council answered that the funding of the judiciary is sufficient to handle their caseload but, unfortunately, it is well known that many Italian courts still suffer a dramatic excessive length of case disposition and a huge number of pending cases, so it is debatable that funding is sufficient. A similar comment can also be made about the Council’s answer that there is sufficient funding to “facilitate judges and other personnel in matters of IT systems, building, etc.”.

Another example comes from question 5e item 3: “Is the promotion of judges is solely based on merit?” The Council institutional and formal answer was “yes”. However, it is well known that the powerful fractions of the Italian Magistrates Association (Associazione Nazionale Magistrati) play a very significant role in judges’ and public prosecutors’ promotions, in particular for the selection of Court Presidents and Chief Prosecutors (Fabri 2016, Di Federico 2012).

This (inter)subjective elements as pointed out by Van Dijk and Vos in their conclusion (p.31) “could make outcomes dependent on the incentives of those who conduct the evaluation”, only partially mitigated by the fact that “the indicators are based on the formal arrangements in a country and otherwise observable phenomena, and can be readily checked by any knowledgeable observer”. I think this can contribute to an improper scorecard assessment of each judiciary.

These examples raise two concerns about the questionnaire submitted. The first one is more specifically related to the formulation of questions, which may lead to different interpretations, and then to improper and a wrong final scores.

21 See for an overview of the questionnaire and the indicators the appendix of the article of Van Dijk and Vos in this Issue.
22 Question 6b is: “Which is the competent body to make the following decisions in the context of disciplinary procedures against judges? It has 6 different items, and for each item the competent body can be “the judiciary”, “the executive”, “the legislature”.
23 For example, this is the case of Italy where in two answers (item d and f) two competent bodies were indicated.
The second concern is broader. Generally speaking, questions posed through the questionnaire focused on the formal legal setting of the judiciary. Therefore, answers do not take into consideration the real practices, which should be the core of the measurement of judicial independence and accountability.

Data collected through the survey submitted to EU judges also raise some concern about the methodology used and therefore the reliability of its outcome. For example, it is not clear who are the “judges” who answered the questionnaire. As the previous section showed, the definition of a judge is quite different in every country and different kind of “adjudicators” may have a different opinion about judicial independence and accountability. As the data show, the number of replies for each country is quite different, in particular considering the various sizes of each judiciary. General “averages” are necessarily affected by these differences, as well as the reliability of some country scorecards.

Some concerns also deal with the relationship between concepts and indicators proposed by the report and by Van Dijk and Vos’s paper. On the one hand, the distinction between “formal requirements” of independence (usually arranged by law), and “perception of independence” (how judges and the citizens perceive their judiciary independent) is convincing. Convincing also is the association between “formal requirements”, and the so-called “de jure” independence, which the authors of the report call “objective independence”. On the other hand, the association between the “perceptions of independence” is much less convincing as a proxy of the “de facto” independence, which the authors call “subjective independence” (Hayo and Voigt 2007). How the formal requirements are met in practice (de facto independence) is of paramount importance as mentioned in Van Dijk and Vos’s introduction (p. 1): “It is important to have a clear understanding not only of what is required for judicial independence and accountability but also to what extent these requirements are met in practice”. At this stage, the ENJC study does not allow to say much on how judicial independence and accountability are dealt with in practice, also due to the methodology used to collect data.

Alternative ways to “measure” de facto independence are briefly presented and then excluded, with good reasons, by Van Dijk and Vos’s paper, p. 9, but still to base the measurement of de facto independence only on perceptions does not match the variable that ENJC wants to measure. Day by day practices have to be observed to check if and how the legal framework is applied in practice. As Van Dijk and Vos point out (p. 25) “In Bulgaria, many of the formal arrangements are state of the art […] There is a strong legal basis for an independent judiciary [however] The perceptions of independence lack behind, which is likely to reflect reality”. On the contrary, (p. 26) countries such as Denmark, Sweden, and Finland with “weak formal arrangements of independence and accountability go together with positive perceptions in society and among judges about independence”. Judges’ and citizens’ perceptions are not good enough proxies to explain how judicial independence and accountability work in practice (de facto).

Perceptions and practices are simply two different things that need to be investigated in different ways.

I share what is written in the ENJC conclusions (p. 31): “As to the way forward, making a systematic assessment of the level of independence and accountability achieved in practice [roman added] by the national legal system is a crucial starting point for improving justice systems across the EU”. This systematic assessment of the level achieved in practice still need some more efforts by the ENJC.

I also have some concerns about the definition and the measurement of accountability. The ENJC report uses the concept of “accountability” and “transparency” as interchangeable. It is understandable that for the sake of simplicity and measurability, the two concepts have been associated, but it is intuitive that transparency is just one part of the broader concept of accountability (Contini and Mohr 2007). Therefore, the current measurement of accountability may consider some more indicators in the future.

5. Concluding remarks

I believe that both parts of this paper show the importance to found any analysis on robust and reliable data, whatever they may be quantitative or qualitative. If data are not reliable, any analysis is undermined in its foundation, and it will be necessary result in misleading outputs and outcomes. This is particularly risky for analyses that should support the policymakers, as argued by ENJC, which may be erroneously induced to design and implement policies based on flawed data. This is even truer for comparative analyses, which are appealing but complex, due to the variety of institutional settings, legal rules, and practices.

As this paper has shown, there are still some severe problems in the comparability of data about the number of judges, court personnel, and court-performance across European countries. CEPEJ is still struggling to improve the consistent interpretation of its questionnaire through a constant revision of the so-called ‘Explanatory note’, and the organization of meetings among the

---

24 Contini and Mohr (2007, p. 30) quote Herbert Simon definition of accountability as: “the combination of methods, procedures and forces determining which values are to be reflected in administrative decisions”. As they argue in their paper, transparency is a pre-requisite of accountability, which has other components such as the correct application of law and procedures, the correct use of resources, a fair treatment of parties.

25 Van Dijk and Vos in section IV of their paper already wider the concept of accountability: “Systematic accountability requires transparency of procedure (e.g., case allocation and complaint procedures) and transparency of performance (e.g., timeliness and efficiency). Individual judges are accountable if their decisions are publicly available, properly reported and explained”.

---
national correspondents. A similar approach, on a limited scale, is also used by the ENCJ. Data collected should be supplemented by more qualitative information to better understand their comparability.

A possible development towards a prudent and progressive comparability of the data across countries could be the creation of clusters of judiciaries, to have groups that are less different in, at least, some of their constitutive features.

Due to the complexity of the different justice systems and their context conditions, the focus of the comparative analysis should be on a limited number of ‘comparable cases’, meaning cases that are less different among themselves (Lijphart 1971). This does not necessarily imply to compare judiciaries that are in a same geographical area because the emphasis is on the features of each judicary and not on their geographical position. The grouping process should adopt a ‘fuzzy logic’, to be flexible in the establishment of comparable clusters. As a starting point, the cluster could include just a few countries and then progressively be enlarged, applying a “fuzzy membership” (Ragin C. and P. Pennings 2005).

The key variables considered to select the ‘unit of analysis’ to be compared across judiciaries have to go through a qualitative and flexible ‘calibration process’ that allow identifying the judiciaries that can be grouped in the same cluster. This process is necessarily qualitative, it uses substantive knowledge, with approximate reasoning to make explicit.

In both exercises, data have been collected through national institutions/correspondents. It could be useful to involve researchers and academics in the validation process of the data collected. In particular, the ENCJ should avoid the risk of insularity in their research activities. There is always a resource problem, but I believe the ENJC’ activities will benefit from the involvement of scholars or other experts, as well as researchers can learn from the practitioners’ perspective.

The ENCJ effort is however; it also suffers from quite a few specific methodological problems. Some of them can probably be solved with some tune-up work (e.g., ambiguity in the formulation of the questions or some pre-set answers, scores coding, involvement of “external” experts), some other will request more resources and much more effort (e.g., sample representativeness, weak connection between the indicators and the variable that has to be measured).

In this regard, the questionnaire filled out by Judicial Councils, or other institutions should be made public for two main reasons. The first one is that they are full of interesting pieces of information that can be very useful for researchers, practitioners, and policymakers. The second one is that this will add transparency to the study, giving the possibility to everyone to assess the reliability of data and analyses. Comments on single countries by researchers or other experts should be welcomed by the ENCJ, since they may improve the quality of the study.

Some more emphasis should be put on the different institutional settings of judiciaries and Judicial Councils, because most probably they affect the scoring and the level of independence and accountability. For example, there are countries such as Italy and Romania where the Judicial Council serves for both judges and public prosecutors. If and how different settings should be considered in the questionnaire, in the scoring, and in the data analysis, are matters of further investigation.

The ENCJ report (p.20), as well as Van Dijk’s and Vos’s paper (p.22), state that “judicial self-government, balanced by accountability, is desirable”. I think this statement is questionable, also based on the data collected with this exercises by the ENCJ. There is no evidence that judicial self-government is per se a guarantee of higher judicial independence and, even less, of higher judicial accountability (Bobek and Kosar 2013, Di Federico 2012). Maybe it is desirable for the sake of judges, but this is a different argument.

This brings me to a more general comment on the usage in the ENCJ report of the terms “justice system”, “judicial system” and “judiciary”. I cannot deal here with this subject in detail, but these terms should be framed better in the ENCJ’s report, because they may lead to some misunderstandings. Some explanation is needed, similar to the one that CEPEJ gives about the budget of the judiciary, in which ‘the judiciary’ refers only to courts.

The ENCJ report (p.20), as well as Van Dijk’s and Vos’s paper (p.22), state that “judicial self-government, balanced by accountability, is desirable”. I think this statement is questionable, also based on the data collected with this exercises by the ENCJ. There is no evidence that judicial self-government is per se a guarantee of higher judicial independence and, even less, of higher judicial accountability (Bobek and Kosar 2013, Di Federico 2012). Maybe it is desirable for the sake of judges, but this is a different argument.

This brings me to a more general comment on the usage in the ENCJ report of the terms “justice system”, “judicial system” and “judiciary”. I cannot deal here with this subject in detail, but these terms should be framed better in the ENCJ’s report, because they may lead to some misunderstandings. Some explanation is needed, similar to the one that CEPEJ gives about the budget of the judiciary, in which ‘the judiciary’ refers only to courts.

26 A similar approach has been followed by the “Court Statistic Project” in the United States (http://www.bjs.gov/index.cfm?ty=dcdetail&iid=283). “The Court Statistics Project (CSP) provides a systematic means to develop a valid, uniform, and complete statistical database that details the operation of state court systems. It provides high-quality, baseline information on state court structure, jurisdiction, reporting practices, and caseload volume and trends [...] The CSP fulfills the vital role of translating diverse state court caseload statistics into a common framework that all states use when establishing their respective goals and policies. Information for the CSP’s national caseload databases comes from published and unpublished sources supplied by state court administrators and appellate court clerks. The CSP has evolved since 1975 by providing more consistent definitions of key terms and parameters for counting. The State Court Model Statistical Dictionary (updated version published in 1989) provided the first set of common terminology, definitions, and usage for reporting appellate and trial court caseloads”.

27 I use here the theory of ‘fuzzy logic’ in a broad sense. It means that clusters of countries should be classified without sharply defined boundaries but based on some constitutive features and the qualitative information collected to establish quite consistent groups of countries for each issue to compare. These cluster groups can vary if the issue to compare varies because the constitutive features of the countries can vary as well. For example, on the number of judges could be possible to create a cluster of 3 or more countries that have several similarities on that issue, but the same 3 countries could not be compared on the number of court personnel.

28 “Comparatists have certainly learned that legal principles are not absolute [...] and the conflict of values has to be reconciled not by the rigor of artificial logic, but by a flexible and pragmatic recognition that [...] a compromise solution has to be formed” (Cappelletti 1983, p. 13).

29 Ragin C., Fuzzy Sets: Calibration Versus Measurement, University of Arizona Paper, www.u.arizona.edu/cragin/fsqca/download/calibration. Download 12 October 2016, p. 1 “Calibration is a necessary and routine research practice in such fields as chemistry, astronomy, and physics. In these and other natural sciences, researchers calibrate their measuring devices and the readings these instruments produce by adjusting them so that they match or conform to dependable known standards”.

---

26 A similar approach has been followed by the “Court Statistic Project” in the United States (http://www.bjs.gov/index.cfm?ty=dcdetail&iid=283). “The Court Statistics Project (CSP) provides a systematic means to develop a valid, uniform, and complete statistical database that details the operation of state court systems. It provides high-quality, baseline information on state court structure, jurisdiction, reporting practices, and caseload volume and trends [...] The CSP fulfills the vital role of translating diverse state court caseload statistics into a common framework that all states use when establishing their respective goals and policies. Information for the CSP’s national caseload databases comes from published and unpublished sources supplied by state court administrators and appellate court clerks. The CSP has evolved since 1975 by providing more consistent definitions of key terms and parameters for counting. The State Court Model Statistical Dictionary (updated version published in 1989) provided the first set of common terminology, definitions, and usage for reporting appellate and trial court caseloads”.

27 I use here the theory of ‘fuzzy logic’ in a broad sense. It means that clusters of countries should be classified without sharply defined boundaries but based on some constitutive features and the qualitative information collected to establish quite consistent groups of countries for each issue to compare. These cluster groups can vary if the issue to compare varies because the constitutive features of the countries can vary as well. For example, on the number of judges could be possible to create a cluster of 3 or more countries that have several similarities on that issue, but the same 3 countries could not be compared on the number of court personnel.

28 “Comparatists have certainly learned that legal principles are not absolute [...] and the conflict of values has to be reconciled not by the rigor of artificial logic, but by a flexible and pragmatic recognition that [...] a compromise solution has to be formed” (Cappelletti 1983, p. 13).

29 Ragin C., Fuzzy Sets: Calibration Versus Measurement, University of Arizona Paper, www.u.arizona.edu/cragin/fsqca/download/calibration. Download 12 October 2016, p. 1 “Calibration is a necessary and routine research practice in such fields as chemistry, astronomy, and physics. In these and other natural sciences, researchers calibrate their measuring devices and the readings these instruments produce by adjusting them so that they match or conform to dependably known standards”.

---
Some more attention should also be addressed to collecting more data on “internal independence” which seems quite neglected in comparison to indicators and sub-indicators used to assess “external independence”.

Moreover, some more information about the sample of respondents to the questionnaire is needed to assess the findings and the score reliability of the survey.

Finally, in some countries more than in others, it would be interesting to carry out a similar but reverse exercise, to analyze the independence of the executive and the legislative from their judiciary and public prosecutor’s office, but this is something that, I suppose, it is not going to be carried out by the ENCJ.

References


Reaction on the comments on the ENCJ study on Method for Assessment of Judicial Independence and Accountability.

Frans van Dijk and Philip Langbroek

We are grateful for the contribution of Stefan Voigt, Elaine Mak, David Kosař, Samuel Spáč, Ingo Keilitz and Marco Fabri to this Special Issue. Their commentaries on the indicators for independence and accountability of the judiciary as developed for the ENCJ give many useful ideas for future development. The comments also reflect the different disciplinary backgrounds of the authors and point to the need to position the ENCJ approach within the diverse disciplines that engage in the analysis of judicial independence. It is obvious that the approaches of the commenters on the ENCJ study differ widely. In economics the approach focuses on measuring independence for inclusion as variable in econometric models about, for instance, economic growth or protection of property rights. More (de-facto) independence enhances economic performance, but how more independence is to be achieved is not addressed. From the perspective of performance management of organizations, independence is part of court performance for the clients and to some degree subservient to it. In a legal, descriptive approach, the situation in different countries is described in detail, also as a part of judicial culture. The ENCJ study only sets criteria for measuring judicial independence, and does not address performance measurement of courts and judges in general.

Independence and autonomy in public administration

At the basis of the approach of the ENCJ is the institutional design of the judicial system. It’s core mission is the delivery of fair trials. Judicial independence is a precondition for conducting fairness in trials. Not the performance of courts and judges, but the constitutional norm that judges must be independent and impartial is at the basis of this study. The study has tried an approach where the current independence is measured and can be used as a part of accountability of judiciaries for the outcome by publishing the outcome.

What does it take for the judiciary to be independent, taking into account the variety of circumstances in Europe? Judiciaries have developed guidelines for judges and maintaining judicial values at a detailed level, as described by Elaine Mak in this issue.1 Regarding the organization and its processes the ENCJ has developed guidelines, for instance about the composition of Councils for the judiciary and proper safeguards for disciplinary procedures against judges. Although the ENCJ’s set of indicators is more normative and has been developed in parallel, its analysis of independence is closely related to that of autonomy in administrative/political science. Autonomy of public agencies in relation to their functioning in complex bureaucracies is a major field of study. Autonomy of agencies and independence of the judiciary have different starting points, where the independence of the judiciary as a whole serves the independence of the judge, but lead to very similar operationalisations. In their review of this literature Verhoest et al.2 (2004) show that autonomy is a multi-dimensional concept, and distinguish managerial, policy, structural, financial, legal and interventional autonomy. These types of autonomy are largely self-explanatory, except perhaps for structural and interventional autonomy. Structural autonomy concerns the influence of government in the decision making body of the agency. Interventional autonomy concerns the possibilities to intervene based on oversight mechanisms, such as performance measurement and auditing. This is the other side of the coin of accountability. The authors note that for an agency to be autonomous, all types of autonomy must be met, and show in their empirical study of public agencies in Belgium/Flanders that this often not the case.

The position of some public agencies is close to that of the judiciary. For instance, the national central banks in the EU are expected to determine monetary policy independently. These national central banks form an interesting case, as at the EU level criteria have been developed for their independence that apply to member states that have adopted the euro or are going to adopt the EU currency (stage three of EMU). This case is interesting not only because of the European framework but also because it was recently developed. In this framework functional, institutional, personal and financial independence are distinguished.3 As to the relationship with accountability the ECB states: “Central bank independence is fully compatible with holding NCBs accountable for

their decisions, which is an important aspect of enhancing confidence in their independent status. This entails transparency and dialogue with third parties.”

The table compares the ENCJ framework for the judiciary as a whole with the framework of Verhoest et al. and for national central banks. It should be stressed that the ENCJ framework also contains indicators about the independence of the individual judge.

Comparison of frameworks for independence and autonomy at organization level

<table>
<thead>
<tr>
<th>Framework Verhoest et al. for agencies</th>
<th>Framework ENCJ for courts</th>
<th>Framework for national central banks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managerial autonomy</td>
<td>Management of court system</td>
<td>Autonomy in staff matters as part of financial autonomy</td>
</tr>
<tr>
<td>Policy autonomy</td>
<td>[Procedural laws vs court regulations]</td>
<td>Functional autonomy</td>
</tr>
<tr>
<td>Structural autonomy</td>
<td>Organizational independence</td>
<td>Institutional independence</td>
</tr>
<tr>
<td>Financial autonomy</td>
<td>Funding</td>
<td>Financial autonomy</td>
</tr>
<tr>
<td>Legal autonomy</td>
<td>Legal basis of independence</td>
<td>Functional autonomy</td>
</tr>
<tr>
<td>Intervenational autonomy</td>
<td>Accountability (inverse)</td>
<td>Accountability (inverse)</td>
</tr>
</tbody>
</table>

We conclude that the ENCJ framework is consistent with conceptual and applied frameworks about autonomy/independence in the public sector. Independence is multi-dimensional and this needs to be reflected in any measurement system. In response to Voigt who states that the available economic empirical research is sufficient for any purpose the ENCJ may have, this literature, while extremely important in its own right, does not address the complex issues at hand when organizing independence. Furthermore, Voigt presupposes that judiciaries and judges are driven by incentives that are economically relevant. Even although European judiciaries are very different in their organisational design, and even although in some countries career-incentives are missing, whilst in others corruption is an issue, these observations should not obscure that in these and other countries most judges show responsible, selfless behaviour in order to serve justice in the cases assigned to them.

Keilitz compares the operationalizations in the ENCJ study with the performance measures in the International Consortium for Court Excellence’s Global Measures of Court Performance, and the NCSC’s CourTools. His criticism of the ENCJ indicators is that they are too many, internally oriented and not accurate enough. He advises to reduce the amount of indicators to a vital few. We think his criticism makes sense from an overall performance of courts and judges perspective, but that goes beyond the scope of the ENCJ study. Judicial independence is a very complex concept and in the end it is related to the judicial conscience. That is what the ENCJ study tries to grasp. Fabri discusses the accuracy of ENCJ’s operationalisations into detail. He welcomes the ENCJ’s efforts against the backgrounds of the methodological weaknesses in the EU Justice Scoreboard and the limitations in the comparability of CEPEJ data. From that position he criticises the ENCJ questionnaire. For example he asks for a clear definition of what a judge is that applies throughout the EU, and he shows deficits in the self-reporting by judicial councils.

Accountability

The operationalization of accountability is criticized by Kosař and Spáč in particular, as missing important issues and including irrelevant issues. Accountability is, obviously, a complex concept, in this context foremost because of its interaction with independence. As clearly summarized by Scholten, independence and accountability are negatively correlated: when greater independence is granted, less options exist for mandatory accountability. This is reflected in the framework of Verhoest et al. by the use of interventional autonomy instead of accountability. If independence of the judiciary is strived for, the instruments for mandatory accountability are limited, and consist mainly of transparency and dialogue, as the ECB notes for the national central banks. This leads to indicators that centre on transparency, explanation and dialogue with society, as in the ENCJ framework. For Kosař and Spáč accountability is foremost about individual judges being accountable to authorities within or even outside the judiciary. Implicitly they presume that judges as civil or public servants are accountable to superiors in a hierarchy. Transparency should include information about disciplinary procedures, but the reliance on discipline to improve the functioning of the judiciary reflects hierarchical thinking instead of promotion of intrinsic motivation of highly trained professionals on which the judiciary depends. The ENCJ indicators in this area cover the protection of judges against the abuse of disciplinary procedures by procedural safeguards. Keilitz describes accountability of judiciaries as a part of the democratic process. Accountability is necessary to inform the general public and policymakers. But measurement of performance in the courts is also necessary for the management of the courts and

---

the judges. In that frame of mind he sees an adequate measuring system of the functioning of courts and judges also as a tool to ward off efforts to politicise the judiciary by ‘Strongmen’. So, while Fabri also calls for an elaboration of indicators of accountability for the functioning of courts and judges, the debate should focus on the question if and how far courts as organisations and judges as professional are or should be hierarchically accountable to superiors. The more they are mandatory accountable, the less judges function independently. Here lies a subject for further research.

Organizational independence and councils for the judiciary
Kosař and Spáč argue that the indicators on organizational independence favour councils for the judiciary, while these councils may be a threat to independence. Voigt suggests a negative relation between having a council and productivity of the judiciary. It should be stressed that the indicator system allows for other governance structures than a council. Furthermore, a council for the judiciary is not a precise concept: councils vary enormously in their responsibilities and in their independence, as reflected in their composition and the way members are appointed. The indicator system solves this by defining organizational (structural) independence as a quasi-continuous variable, mainly consisting of the scores on two factors: the breadth of responsibilities of a council and adherence to the minimum standards of the ENCJ as to its independence. In reply to Voigt, it shows that many councils have no competences to influence the efficiency of the judiciary and these cannot be seen as the cause of inefficiency.

In his recent review of the agency literature Verhoest sees a shift of research focus to the dynamic nature of agency autonomy and control and to the relationship with the increasingly multi-level administration within the EU (Verhoest 2017). Both aspects are observable for the governance of the judiciaries within the EU. For instance, the demarcation of the required independence of the judiciary and, in particular, a council for the judiciary, is dynamically developing. The ENCJ has developed guidelines that recently led it to suspend the Polish Council for the judiciary. The ENCJ is just one of organizations/networks active at the EU level, and its guidelines are, at best, soft law, but these may help to gradually develop hard law. The EC and the CJEU play crucial roles in this process. This gradual emergence of standards would offer opportunities to make the indicator system less subjective than, as pointed out by the commentators, it can now be.

Conclusion
The ENCJ framework of judicial independence and accountability is consistent with frameworks developed in parallel for other parts of the public sector. In particular, the multi-dimensional nature of independence is crucial to recognize. For instance, in many judiciaries, also where the other dimensions of independence are well take care of, lack of financial independence is seen to undermine the functioning of the courts. This basic structure of the indicator system is, therefore, in our view sound. However, there is much room for improvement, for instance, by the inclusion of hard data, extension of surveys to the clients of the courts and refinement and, where possible, simplification of the indicators. At this stage, the normative and subjective nature of indicators is unmistakeable. It is likely that by the gradual emergence of EU-standards the subjectivity will decline over time. In that context, more research on the development of better indicators for the evaluation of performances of courts and judges is necessary!

8  Kosař and Spáč, supra footnote 6, pp. 41 and 43.