2498-5473/USD 20.00
© 2019 Akadémiai Kiadó, Budapest

HUNGARIAN JOURNAL OF LEGAL STUDIES 60, No 1, pp. 119–126 (2019)
DOI: 10.1556/2052.2019.60108

Mátyás Bencze, Gar Yein Ng (eds.): *How to Measure the Quality of Judicial Reasoning.* Springer, 2018

LEA PŐDÖR*

'Measuring the unmeasurable' – this is the main concept of the book and it is also the title of the editors' introductory study. This phrase also wonderfully describes the research project's and the authors' hard challenge. To write about something and then to explain its subsequent quality, judicial reasoning in this case is something very unique, individual, complex and different on a case by case basis. The book is edited by Mátyás Bencze and Gar Yein Ng and the volume concentrates on the quality of judicial forums' reasoning method. The project was based on the cooperation between Hungarian Academy of Sciences, Centre of Social Sciences, Institute for Legal Studies and the University of Debrecen, Faculty of Law. The book consists of fifteen studies, conducted with the assistance of several prestigious legal scholars and legal practitioners. The book has two main parts, in which the first contains essays from the perspective of legal theory, legal sociology, philosophy and behavioural law and economics. The other part colligates the main conclusions from the practice of some national and international judicial forums.

Judges, judicial decisions and judicial argumentation have always had both academic and societal attention, so it may be thought that the issue is well and widely known. It is an age-old experience that judges have three important challenges: decide every single case they get; decision should be based on written rules; and (last but not least) decision should be fair and just.[1] But what is praxis's reality, would judicial work be this simple or not? It is well advised to think over the problem and to concentrate on the quality of reasoning which has become an accentuated requirement in the past years.

The purpose of this research project was to properly answer a difficult question – what is expected from judicial reasoning? Are there any aspects, standards, or scales, which can highlight and measure the quality of judges' work? There are some measuring methods from the legal philosophy tradition, e.g., there are great maxims in Roman law, such as *ius est ars boni et aequi, summum ius summa iniuria,* etc. Legal culture has also created several famous role-models in connection with judges' reasoning methods; these famous descriptions consist of judges who represent the main features of an ideal decision-maker.[2]

* PhD student, Széchenyi István University, leapodor@gmail.com.

[1] Coing (1996) 258.

[2] We should imagine a scale with two widely different endpoints: Weber's judge and the Swiss Civil Code's solution to application of law. Weber expresses that the ideal judge works like a paragraph-machine: clients have to throw the facts of the case, the rules and the legal expenses into this 'machine', which then delivers the judgement. The Swiss Civil Code's solution is interesting, too: if there is not any rule or custom that could decide the case, judges should be legislators. So, between these two endpoints, there are a lot of types of judges. There are two more examples which cannot be ignored: either of them is Dworkin's Hercules-judge who has superhuman skills and according to the scholar, Hercules can solve the so called hard cases as well. The other one, Paul Magnaud is labelled as 'the good judge' – he willingly returned a verdict *contra legem* because he tried to be more fair than written rules had suggested.

Nowadays, a natural expectation can be noticed in real life as well as in the world of science. Trying to understand a dilemma, things get easier if there are numbers, scales, etc. which help form an explanation to a phenomena. Overall, if there are the objective measuring aspects, solving the problem is like going along a straight path – it is safer and more acceptable than any other alternatives. But what if a purely intellectual activity (adjudication) is wanted to be measured? Is it possible at all? Quality of decision-making has always been essential because legal systems' 'operators' and 'users' wish to have legitimate and justified judgments. This project has demonstrated several standards that judgments should at least match, and which may represent the nature of the assessment of judicial reasoning.

Zenon Bankowski's study (25–41) clarifies a theoretical pre-question: what is the expected behaviour and role of a judge. Generally, it is an eternal question of every epoch and the essay starts with its symbolic character, the 'Unjust Judge' who decides wrongfully from the famous parable about the unjust judge and the persistent widow from Luke's gospel.[3] Firstly, he does not fear God and secondly, has no regard for man. Thanks to the widow's persistence, the judge grants her justice – but 'for thought we might fear God and know our doctrine, we might still not know enough to fit the widow uniquely into the law.' (26) The 'Anxious Judge' is aware of something important – knowing the rule is never enough and rules cannot be automatically applied as every case is particular. The 'Judge in the Middle' expresses the judges' space that 'this place can be seen as standing between universalism and particularity; between the rationality of law and the contingent other or between the law and the existential encounter, which is its application.' (27) The author explains some more situations (so-called robojudges, the 'Kantian Saint', and what if hard cases make bad law, etc.), but the main question remains very intense – how to cope with the particularity of cases? The key of understanding can be reached in a special phenomenon – practical intelligence. There is a space ('the middle') between the universal law and the particular decision,[4] and there are some skills and capacities which may help to operate at the limits of language. Of course, law is a text-based discipline but judges should let the case speak for itself.

The social context of measuring quality is analysed by Zoltán Fleck (43–55). It is a general experience in rule of law systems that a wide gap exists between positive law and justice – the legally good is not always ethically right and laymen's feelings and thoughts about justice are far away from professional logic. It is a relevant factor that ethical expectations of society are essential in modern democracies. The measurement of judicial reasoning has to be qualitative in nature as it is also the only way to do a respectful analysis with societal needs and constitutional values as well. Moreover, if judicial argumentation reflects moral values, the assessment system has to take them into consideration. Then, how can moral compliance be measured? It is sure that high-status legal documents always include moral values, so using this type of sources can be a primary test. It should be declared again that the measurement of judicial argumentation has to be qualitative in nature, because this is the only way to make the viewpoint capable to be open to societal needs. In this, lies a paradox – any kind of professional control discussed is always a part of the wider social control. Therefore, if the ideology and public values are in contrast with

[3] See Luke 18, 2– 8.

[4] The coherence between the universal law and the particular decision is well-known because it was greatly analysed by Hart. He expressed his view with two terms: core of meaning and core of penumbra. Hart (1995) 147–59.

the rule of law, qualifying the judiciary will be a dead end (46). It is important to know that measuring means administering at the same time and judges have to control and comment on the system of measurement as well. If understanding what does 'quality' means is really desired, the answer may show the characteristics of the authority which measures it. While measuring quality, it tends to be sensitive as measuring is in connection with the level of independence and the status of judicial autonomy. No doubt that whatever type of measurement is used, it is acceptable if the aim is legitimate. What is the main conclusion? 'We should take the wider context of accountability for measuring quality into consideration. The revolution in accountability and the necessary improvement of judgeship for the sake of citizens' rights could become a disaster if the overt and hidden functions of hierarchical control are not compatible with the basic mission of judicial power.' (54)

Mariusz Jerzy Golecki's aim (57–72) is to elaborate an interdisciplinary model of adjudication, which includes aspects from cognitive psychology, cognitive sciences and behavioural law and economics. Why is it so important to think over the conventional role model(s) of adjudication? These days, there is a curious situation being revealed as judges are faced with affects, heuristics, biases and manipulations. According to the author, the objects under analysis are judicial rationality and impartiality. Law and economics scholars have invented a grand theory called 'rational choice theory' for rationality of decision-making process that is based upon stability, durability, constancy and persistence. Such rationality is a good basis to predict future actions and events. Bounded rationality is a well-known phenomenon in cognitive psychology and has also been applied in economics and legal theory. This concept reminds everyone that human agents, as well as judges, have limited computational skills and memory and unfortunately, they always have limited information.[5] In connection with this information, dual process theory (DPT) can help a lot to understand judicial rationality. The theory consists of an intuitive PT1 (the process is unconscious, automatic and fast) system and a deliberative PT2 (which is awareness, slow and has limited capacity) system. The author also demonstrates the problems linked to the influence of intuition, heuristics and biases. He is also interested in the effects of emotions and highlights the importance of the 'intuitive-override model of judicial decision making' (IOM, which can be a powerful extension of the DPT). The model presents the following process – a judge makes their intuitive decision under System 1 and then compares it with the results of the deliberative and conscious System 2. This is acceptable because firstly, decision tends to be intuitive but, on the other hand, System 2 helps to find rational and valid reasons rather than feelings. It seems judicial decision-making is like a battle of ration and emotions, and it is sure that these two have a close and complicated relationship. At this point, it is useful to cite a thought as Golecki did: 'intuition can be surprisingly accurate, but sometimes good judgment will require purging the deliberative processes of intuition's influence. (…) Furthermore, some decisions might require shifting between both of the systems.' (68)

It is evident that lay participation is a prevailing institute in several states of the world.[6] The cardinal question in connection with this old and dignified apparatus that what are the advantages and disadvantages of applying laymen in judicial decision-making process and why modern legal systems support or refuse them? In the 21th century, most of the states

[5] Bix (2009) 26.

[6] Regarding the European Union, lay participation is quite popular as only the Netherlands and Latvia are the exceptions where laymen are not occupied in jurisdiction. Badó (2017) 99.

conceive that lay participation subserves justice to be better, more democratic and more efficient. Attila Badó analyses (73–86) four aspects of this topic: self-governance, strengthening the legitimacy of justice, control function and lay persons as fact finders. Democracy has always been an important value as it expresses that citizens can take part in decision-making. Numerous theorists and jurists think that laymen, especially jurors, could be the keystone of democracy, a guarantee of freedom. There are opinions believing that lay participation contributes to the legitimacy of the political system. Surveys based on empirical examinations greatly demonstrate that trust in justice has increased thanks to presence of lay bodies. Moreover, there has been a measurable increase amongst those people who tend to participate in deciding legal disputes. An extraordinary myth exists that lay persons could have acumen and could be more vigilant in observing the facts than professional judges. However, there are some special actions in fact-finding which can only be applied by professionals. Of course, if a judge is tired of life-long work and education, fact-finding can be accomplished by anyone, without any erudition. Regarding the control function, it is certain that lay persons bring the so-called 'popular sense of natural justice' into the legal system and can also stand for the social sense of equity. Moreover, lay bodies can represent the humanistic aspects over the bureaucratic justice system. So, all in all, the demonstrated arguments show reciprocity with the image of quality.

Justification is the most relevant indicator of the judgment. The public character of reasoning can defend judicial forums from political pressure and from other powerful social groups, but if pressure like this has any effect on the decision, it may result in poor quality reasoning. The study's author, Mátyás Bencze has three questions (87–101). Firstly, it should be made clear that it is good to have some general requirements to test the quality of judicial reasoning. While elaborating these, attention should also be made to cultural differences, because legal systems have many different features. However, it is necessary to have some universal requirements. Who measures? Internal and external evaluations can come into question with the former, meaning judges, evaluating colleagues. This is problematic with respect to arguments in connection with subjectivity, collegial bias, conformity. The latter is also a sensitive topic, according to the author, as the danger is a potential violation of judicial independence. To sum up, the person who has competence to evaluate can be anybody affected by the activity of the court, and on the other hand, legal scholars, researchers, journalists, etc. The most difficult task is to answer the problem – how could be the quality of judicial reasoning quantified? Monitoring adjudication means human skills and of quantifying the 'signs of good in a product of the human mind.' (96) There are two possible ways to qualify and both can be expressed numerically. The direct way reflects on the generally used legal arguments with the understanding that a greater number of arguments used by judges will result in a greater sensitivity to the reasons behind cases for judicial practice. The indirect form '(…) can infer the quality of legal argumentation from an examination of the reactions to and reflections on the judicial reasoning delivered in the legal system we are examining'. (97) The author then demonstrates some examples in connection with the indirect evaluation.

The book's second part involves studies about some national and international judicial forums' working process.

Gar Yein Ng demonstrates (103–121) that in England and Wales, 'quality-management' of judicial forums is seems to be a well-established process. This does not mean that managerial principles would be the basis of quality of judicial reasoning rather it is founded on the definition of the role of the judge. The judge should stand for the correctness of the decision; pay regard to relevant facts and respect the expected high standards.

The literature contains two questions and their proper answers – why and how should reasons be given.[7] The author refers to 'the pull of justification' which means that judges should always evoke reasons, it is an important obligation. Generally, legal reasoning is based on interpretation and a lot of methods exist. 'If the judge is clear on why she (he) is giving reasons, the how, based on the models (…), should become clear.' (110) In connection with this realization, desired skills and knowledge call for explanation, can be found from official frameworks by the Judicial Studies Board. System of appeal can control the quality of judicial reasoning but the author cites some more criteria from the literature and from academic area in which quality appears. The most important judicial obligation in England and in Wales is giving adequate reasons while drafting the judgment.

Focus changes to Italy with Francesco Contini's survey (123–140) showing that the question of quality is connected with timeliness and the limited available resources. There is a backlog which unfortunately could affect the judicial administration and this is why amelioration of quality is so accentuated. To cope with the problem, it is advisable to take care for increasing efficiency and timeliness. Moreover, there are some 'alarm-words' which may represent the searched quality: legal certainty, treatment of the parties and procedural efficiency. At this point, reference is made to an interesting feature in the world of judicial forums. In this country, a unique scoring system exists to ameliorate judicial work using the following system: A judge receives twenty decisions of another judge, which should be evaluated using several aspects. A negative evaluation results in this judge undergo training and further assessment. If at the end of this process, a judge still fails after two years, they will be dismissed.

The problem of quality is also complex in France and as Emmanuel Jeuland writes (141–154), the quality of justice is not really assessed, moreover much criticism can be found. There are some numerical analyses helping to elaborate the productivity of French courts. In 2011, new criteria have been elaborated to outline what makes judicial work more optimal and effective. In France, judges have to aspire to make a maximum number of judgments – as quickly as possible. It is a bit dangerous as it is known that a decision of good quality demand time! Experts have established that French judges should practice with a greater motivation. There is a fertile desire in every judicial agent as they prefer to have more predictable decisions because a lot of different answers exist for the similar legal questions, making things more dubious. Consequently, it should be determined to evolve a fruitful dialogue between judges. The author also expresses it is well known in France that the question on quality cannot be unified. There is need of more new indicators that could 'reintroduce quantity into quality'. (153) Some of these include harmonised case law; organizing the return of appeal judgments for first instant judges; improving relationship between courts; more efficient education of lay judges; reasoning ignoring stereotypes and looking forward the criticism of the academic community.

'Quality control has taken place very much through the discussions between courts and legal scholars' (155) in Finland. Markuu Kiikeri's study explains the achievements of the main Finnish quality project called Rovaniemi. In Finland, situation of tribunals is quite encouraging because there have been some project in connection with improving judicial work. One of the most important projects is the Rovaniemi Appeal Court. The main goal

[7] In connection with giving reasons, it is also important to clarify why the so-called argumentative method has become an accentuated process. It took a long time to realize, but questions do not have only one proper answer. Therefore, if there are a lot of convincing and rational reasons, there would be more and more „proper" answers as well. Szabó (2015) 211.

was to change the legal environment of courts. The working group established the aspects of the examination in four categories: procedure, decision, client service and organization, and at the end, in a study, seven criteria have been phrased for the quality of decisions: legality, motivations, publicity, analytical quality and consistency, linguistic clarity, structural comprehensibility and clear announcement of the decisions. Moreover, the Supreme Administrative Court has also a report from 1996 concentrating on arguments and good reasoning. The author cites some important paragraphs from the Finnish law to represent the autonomy of the judiciary and to make clear that the constitution declares motivated decisions for everyone. Furthermore, legal theoretical viewpoints of decisions and quality assessments should be known better – this is the task for legal scholars.

In Czech Republic, there is no model for judges on how to make ideal reasoning to their decisions. What is more, quality assessment is in conflict with the separation of powers and judicial independence. This makes things more difficult, because with the existence of an accepted method, judiciary would be united. According to Zdenek Kühn (173–186), the question is how can a judge act where there are several styles and solutions for reasoning. Amongst these, the legalistic style of the Supreme Court and the dialogical and discursive of the Supreme Administrative Court should be separated. It would be a bit rash if just one was stated for convenience – more time is needed to accept a viewpoint. In this situation, one solution is given: 'whether a longer decisions with more reasons is more persuasive within a civilian legal culture, or whether the judiciary by doing this is not losing part of its (…) legitimacy as an institution endowed with the knowledge of the 'objective' and 'one right answer' in the interpretation of the law.' (185)

Last but not least, there are two significant and official solutions used by the NOJ[8] on how to check the quality of justification in Hungary and this was analysed by Mátyás Bencze, Ágnes Kovács and Zsolt Ződi (187–205). One of these solutions is called appeal ratio, which has proved to be a relatively reliable one as it is only based on a court 'user' assessment. If the parties are satisfied with the decision, it will show the lack of appeals, which could also indicate the judgment's convincing power and the high quality of reasoning. There is an available comparative data demonstrating civil cases from 2011 and it helps to reveal interesting tendencies on the prevalence of appeals. What is more relevant here, is there a special correspondence: 'the variety in the appeal ratio can be a sign of variety in the quality of justification. (…) the satisfaction level of the parties correlates (…) their perception of how the court treats them.' (190) Personal assessment is the other method and is performed by individual judges – the immediate professional superior, who knows the judges personally. Judges have to go under examination in the third and sixth years from their appointment and then every following eight years. The assessment's aim is to evaluate the quality of judge's work but here is a problem that the assessor and the assessed judge are from the same county. Furthermore, in Hungary, a style-guide is achievable for judicial forums and is based on a report created by a drafting group and summarizes improvement proposals in four fields: grammatical and stylistic level, unifying the writing of certain words, unifying the structure of the decisions and proposals for the substance of the reasoning. The study exposes the greatest remarks of the report, besides the fact that the proposals were formally accepted, the changes were only introduced at the Curia. It is good to know that remarks of a style guide cannot change the culture of the judiciary, and such requirements can always be formalistic and impersonal. To clarify

[8] National Office for the Judiciary.

more additional indicators, European Court of Human Rights (ECtHR) and Hungarian Constitutional Court (HCC) have an effect on legal argumentation of Hungarian courts as both of them control the way of Hungarian courts' justification. To be honest, it is a bit doubtful whether human rights tribunals can be indicators of the quality of justification…

Marjan Ajevski introduces (207–224) the situation of the International Criminal Tribunals. The author's task is difficult as he admits truthfully that these types of courts are a slightly odd amongst the 'traditional' ones. International criminal law is also a special area in the world of law, moreover the procedure has been formatted by judicial discretion. This is the extraordinary condition where the establishment of the International Criminal Court (ICC) successfully changed the fractured landscape of international criminal law. All in all, this landscape's eternal featuring is a bit *ad hoc,* but in the past twenty years, ICC has its own mechanisms which help to deal with criminal cases. Ajevski has an idea that it is a bit early to draw a sketch on the quality of reasoning at the ICC but there are some interesting factors in the background. ICC can discuss the format of judgments on a formal (title, parties, summary, details) and on a substantive (reasoning) level. The length of the proceeding is a critical point as a process usually last for years (of course, experts have revealed lots of arguments, for example the great number of participants). What is more, the length of judgments is flabbergasting, usually several hundred pages long. The decisions involve the majority and the minority's opinion (separate and dissenting opinions) as this method belongs to the international tradition of international law. To make things more interesting, that practice has consequences regarding the length of the judgment and the timeliness of delivery. The ICC strives to make the judgments accessible so decisions are made in the official languages of EU but usage of a different language is allowed. Maybe the most relevant question may be the rules of interpretation. Nowadays, ICC is in a lucky position as it has a more detailed statute and the hierarchy of laws is already clear. The actual method in interpretation is based on good faith, and the statute establishes the single rule of interpretation with only three elements: ordinary meaning, context, object and purpose (224).

The essential information is that the expression 'the quality of decision-making' has occasionally appeared in the literature on the Court of Justice of the European Union. Quality is in connection with legitimacy, which relates to the values of the rule of law, democracy and human rights. At first, regarding the mentioned quality, Gerard Conway (225–250) makes an effort to illustrate legal methods in different aspects, which could be the criteria of quality assessment: normative/institutional legitimacy; epistemic quality; argumentative, dialectical or deliberative quality; consistence/coherence; moral and consequentialist quality and last but not least, jurisdictional quality. Characterising the quality is reliant on the conception of quality employed, and it can be expressed that quality may have success. This helps the discussion about the measurement. CJEU has a practical success as judgments are usually accepted (and of course applied) by the member states but it is not lucky to draw conclusions on the decisions' perfectness e.g., dogmatics and methodology. The author's main conclusion must be cited here. 'Quality (…) is to be judged by the Court's adherence to and articulation of the conventional norms of legal reasoning focusing on the ordinary wording of the most specific texts, drafting history as evidence of the intention of the law-maker, and an attitude of deference to the constituent power of the formal process of Treaty amendment.' (247)

Because of the fact that the previous section concentrates on jurisdiction of states of European continent, one more coherent issue has remained to examine: case law of the European Court of Human Rights, where there are requirements concerning quality

of reasoning. David Thor Björgvinsson opposes (251–268) that ECtHR is in a merely inimical environment; what is more, the forum should face legitimacy issues and has already survived some judicial restraint, too. Björgvinsson seems to be critical when he argues regarding reasoning, it is not sure there will be logical judgments – but uncertainty and consistency are fundamental values. In the study, there is a landscape of the traditional canons of interpretations (intentionalism and textualism), which aspect has coherence with judicial restraint/activism. If consistency is not respected, people will lost their confidence in the judicial system. It is in question whether the famous concepts – subsidiarity, margin of appreciation and consensus – could serve as methods of interpretation. It is apparent that the ECtHR has changed direction as it abandoned its dynamic approach, and as to quality of reasoning, there is a solution which could be a compromised one.

Reading this well-founded and high-quality volume, we must admit: 'measuring the unmeasurable' is not so impossible. The editors' introductory study refers to a meaningful motto: 'If you can't measure it, you can't improve it.' (2) In a sense, it was a bit ambiguous whether judicial reasoning can be measured, but there is some certainty that the quality of judicial adjudication does matter. To admit, a lot of circumstances must be discovered to have a transparent landscape for the analysis. Regarding the mentioned studies, most of the methods contain efficiency issues and of course, the research cannot be separated from adjudication's special social and institutional background. To sum up the main and common idea of the authors, it is better to once more cite the editors' thought: '(…) there is a silent consensus amongst our authors that from the outside of the judiciary only 'soft' methods such as public and academic debates, as well as published good practices can be incentives to improve the quality of judicial reasoning. This appears to answer the question of whether the idea of the 'quality control of judicial reasoning' is any given legal system.' (22)

LITERATURE

Badó, Attila, *A bírói függetlenség egyes garanciális elemeinek összehasonlító vizsgálata* (The Comparative Analysis of Certain Elements of Judicial Independence) (Dsc. Dissertation 2017).
Bix, Brian H., *A Dictionary of Legal Theory* (Oxford University Press 2009).
Coing, Helmut, *A jogfilozófia alapjai* (The Basics of Legal Philosophy) (Osiris 1996).
Hart, Herbert L. A., *A jog fogalma* (The Concept of Law) (Osiris 1995).
Szabó, Miklós, *Rendszeres jogelmélet* (Legal Theory in System) (Bíbor 2015).