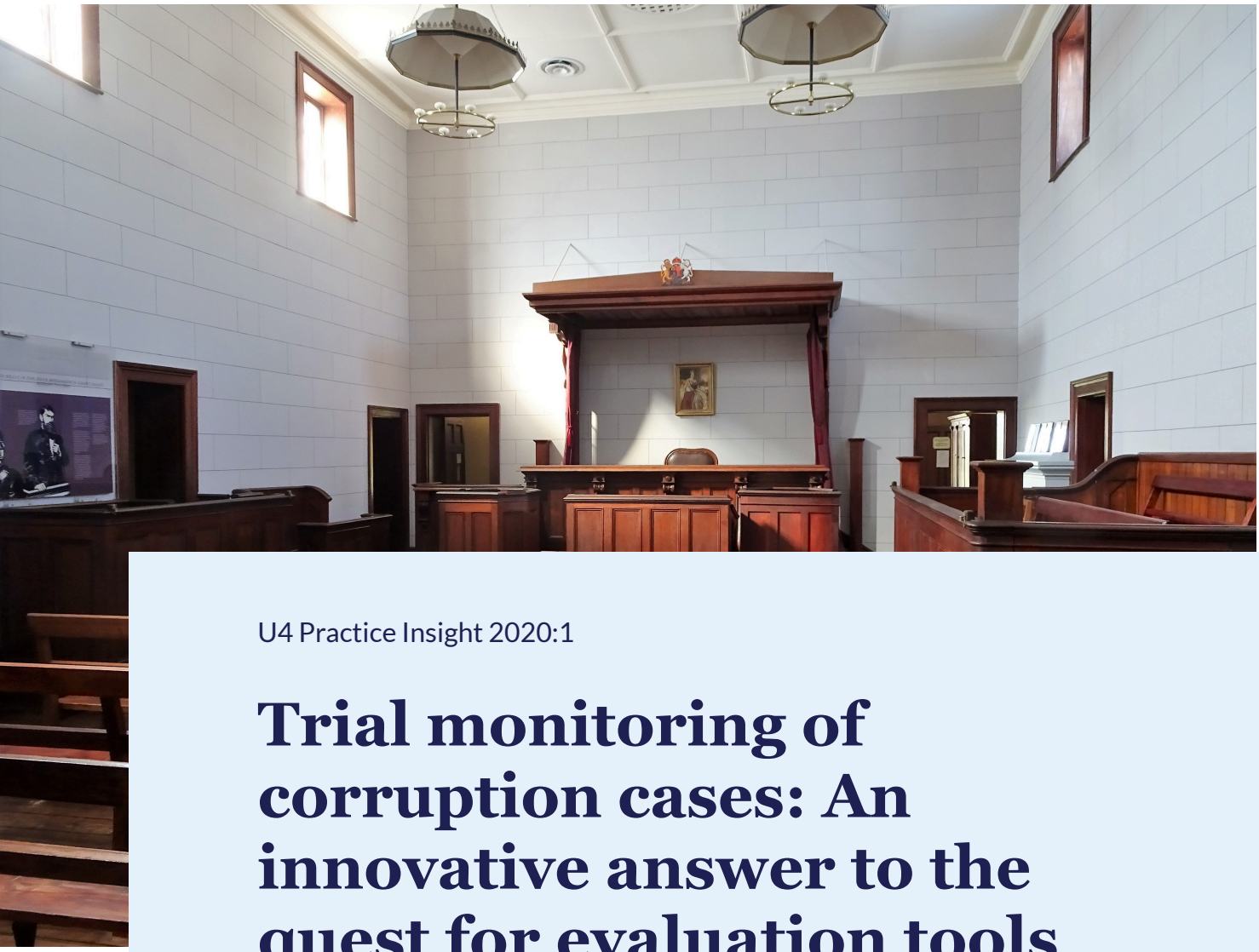
U4 ANTI-CORRUPTION RESOURCE CENTRE



U4 Practice Insight 2020:1

Trial monitoring of corruption cases: An innovative answer to the quest for evaluation tools

By Francesco De Sanctis
Series editor: Sofie Arjon Schütte

CMI CHR. MICHELSEN INSTITUTE

Disclaimer
All views in this text are the author(s)', and may differ from the U4 partner agencies' policies.

Partner agencies
Australian Government – Department for Foreign Affairs and Trade – DFAT
German Corporation for International Cooperation – GIZ
German Federal Ministry for Economic Cooperation and Development – BMZ
Global Affairs Canada
Ministry for Foreign Affairs of Finland
Ministry of Foreign Affairs of Denmark / Danish International Development Assistance – Danida
Swedish International Development Cooperation Agency – Sida
Swiss Agency for Development and Cooperation – SDC
The Norwegian Agency for Development Cooperation – Norad
UK Aid – Department for International Development

About U4
U4 is a team of anti-corruption advisers working to share research and evidence to help international development actors get sustainable results. The work involves dialogue, publications, online training, workshops, helpdesk, and innovation. U4 is a permanent centre at the Chr. Michelsen Institute (CMI) in Norway. CMI is a non-profit, multi-disciplinary research institute with social scientists specialising in development studies.
www.U4.no
U4@cmi.no

Cover photo
denisbin on flickr (CC by-nd) https://flic.kr/p/Hpx3Yz

Keywords
Bosnia and Herzegovina - Southern Europe - justice sector - judiciary - courts - trial - monitoring, evaluation, and learning

Publication type
U4 Practice Insight

Notes
Watch: Video interview with Francesco De Sanctis sharing insights on trial monitoring

Impartial, competent and efficient judiciaries are vital in addressing corruption. Both qualitative and quantitative aspects of performance should be considered when assessing the effectiveness of a judicial system in handling corruption cases. Trial monitoring can substantially enrich the evaluation and analysis of the judicial response to corruption.

Main points

- The prosecution and adjudication of grand and political corruption present unique challenges due to the legal and factual complexity of these crimes and the risk of undue political pressure in the process.
- Minimal attention has been given to devising and applying criteria and instruments for assessing the performance of the judiciary in processing corruption cases. In order to develop satisfactory tools for this task, it is important to understand those practices, policies, and reforms that work, and those that do not.
- Trial monitoring programmes have the potential to strengthen how we evaluate judicial response to corruption. A successful programme should incorporate a methodology for assessing all the features that a judicial system should have – impartiality, competence, and efficiency – and both qualitative and quantitative aspects of performance should be considered when assessing the effectiveness of a judicial system.
- Trial monitoring by independent organisations can be used as an effective instrument to enhance public and external scrutiny on the work of the justice system in tackling corruption.
- A new approach to trial monitoring could be adopted and implemented by international actors active in the field of anti-corruption and judicial reforms, and would enable the comparison of results across different jurisdictions.

Table of contents

About the author

Francesco De Sanctis

Francesco De Sanctis is a lawyer with fifteen years of professional experience in the areas of anti-corruption, criminal justice reform, and transitional justice. He is currently working as an expert for the OSCE Mission in Bosnia and Herzegovina, where he previously held the position of Head of Rule of Law. In the past, he worked as a consultant in several internationally funded projects and for the EU Rule of Law Mission in Kosovo (EULEX) as legal officer. He has a Ph.D and an LL.M in international human rights law and published articles on international law reviews, including the International Journal of Criminal Justice. De Sanctis can be contacted on fradesan@gmail.com.

Acknowledgements

U4 and the author would like to thank Rosanna Amato, Francesco Contini, and Drino Galicic for their insightful and constructive comments on an earlier version of this paper. The author would also like to dedicate a very special thanks to his wife Maja Pulic for her patience in introducing him to social research methods and concepts which were foreign to his legal background.

The role of the judiciary in addressing corruption is fundamental.[1] However, the prosecution of grand and political corruption presents unique challenges due to the legal and factual complexity of these crimes and the risk of undue political pressure in the process.[2]

Countries in different parts of the world are investing political capital and trying different institutional settings to tackle these problems. An example of these efforts is the creation of specialised prosecutors' offices and/or specialised courts to process corruption cases.[3] Even where judicial bodies with general jurisdiction are in charge, the drive towards specialisation in this field can be observed just by browsing the numerous manuals and training curricula available on the internet which are devoted to the investigation, prosecution, and adjudication of corruption cases.

Against the background of these developments, little attention has been paid to devising and applying criteria and instruments for assessing the performance of the judiciary in processing corruption cases. The importance of such endeavour should be evident. It is not possible to understand which practices, policies, and reforms work (and which do not) without adequate assessment tools.

Evaluating judicial response in the processing of corruption cases, however, is not a straightforward exercise. It involves a rigorous process composed of distinct steps. After defining the topic, the evaluators need to identify the research questions (in this case: which questions about judicial performance relating to the processing of corruption they intended to answer). Then they need to choose a methodology (based on quantitative or qualitative methods, or a combination of both). Only at this point can the selection and collection of the relevant data start, which will result in their analysis and presentation of the results.[4]

I went through this process, and its many underlying questions, when I started working for the OSCE Mission to Bosnia-Herzegovina ('the Mission') on a project named 'Assessing Needs of Judicial Response to Corruption through Monitoring of Criminal Cases' (ARC). I was in charge

1. Ferguson (2017). Please note that for the purposes of this paper the term 'judiciary' includes both judges and prosecutors.
2. ADB-OECD (2013).
3. Stephenson and Schütte (2016).
4. Simion (2016).

of devising a trial monitoring-based methodology which would be suitable for analysing problems and trends in the effectiveness of judicial response to corruption, and proposing measures aimed at strengthening the capacity of the judiciary in processing corruption cases.

Based on this methodology, the Mission, through its established trial monitoring programme, gathered the data that were analysed in two reports (published in February 2018 and April 2019) assessing judicial response to corruption in Bosnia and Herzogovina (BiH). As the two reports are fact-finding oriented and aimed at informing judicial and anti-corruption reforms in BiH, they make only brief reference to the methodological tenets of the project. For this reason the reports allow only limited insight for actors, institutions, or organisations which would be interested in replicating this methodology to monitor other domestic judicial systems.

This paper offers a practical contribution to the topic of evaluation of judicial performance in the processing of corruption cases. I will present the methodology devised for the project, explain how it was applied in the context of BiH, and argue that it could be effectively employed in other countries or regions. The value of this effort lies in my proposition that the methodology in question is innovative to the extent that it combines, in an original and purposeful manner, two sets of instruments which have so far been used in distinct and unrelated efforts. There are those that are devised by state institutions to assess the performance of the justice system as a whole or of judges and prosecutors individually; and those employed by international organisations and civil society organisations (CSOs) in the context of trial monitoring activities, mainly for the purpose of assessing the compliance of the justice system with international standards.

Defining effective judicial response to corruption cases

Judicial institutions can be assessed according to a number of different methods and criteria depending on the aspects of their performance that are considered and the analytical tools used. In their overview of evaluation methods in different European countries, Mohr and Contini have identified various approaches. Virtually all evaluation systems propose, as a minimum, an assessment of the efficiency of judges and prosecutors, mainly in terms

of productivity and timeliness of proceedings, through the use of quantitative indicators.[5]

In addition to quantitative aspects, many systems evaluate the quality of justice that is delivered. This may take different forms that can be traced back to two different notions of quality. In a stricter sense, it is concerned with the quality of judicial and prosecutorial decisions; in a broader sense it encompasses the quality of the judicial process and its management by considering aspects such as accessibility, treatment of parties, impartiality, and fairness.[6]

Before considering whether institutional evaluation systems can help us to assess judicial response to corruption, we need to determine which aspects of judicial performance are relevant for this endeavour. Corruption cases are rather specific in nature, so it would seem logical that an assessment focussing on the processing of these cases should be tailored on that specificity. They are often more complicated than standard cases (as they may require particular expertise), and more likely to be the object of political pressures and interferences, especially when high-ranking or influential subjects are involved. As a result, it can be argued that, in a well-tailored evaluation method, consideration of competence, independence, and impartiality of the judicial actors should be prominent.[7]

A well-tailored evaluation method, consideration of competence, independence, and impartiality of the judicial actors should be prominent.

An assessment of these qualitative features, albeit fundamental, would give us only an incomplete picture if not accompanied by proper consideration of relevant quantitative aspects. Issues of performance related to efficiency, in terms of both productivity and timeliness, must be properly assessed. It is important to emphasise that quality and quantity of judicial performance should not be seen as contrasting goals. Although there is a real tension

5. Bencze and Ng (2018).
6. Albers (n.d.).
7. Stephenson and Schütte (2016).

between the two, they are strongly interlinked. A judicial system that impartially and correctly adjudicates on appropriate charges, but processes only a few minor cases or is affected by undue length of proceedings, could not be said to fulfil the general goals of criminal justice.

Therefore, I find that the concept of effectiveness best encompasses all the features (independence/impartiality, competence, and efficiency) that a judicial system should have, both generally and with regard to the processing of corruption. While it may be difficult to define precisely what effectiveness is in the context of criminal justice, it indicates whether or not the institutions in question are achieving the goals that society has set for them.[8] The general goal of the criminal justice system is to deliver justice by convicting and punishing the guilty while protecting the innocents.

Similarly, the World Justice Project (WJP) defineseffective criminal justice systems, for the purposes of its Rule of Law Index, as systems that 'are capable of investigating and adjudicating criminal offenses successfully and in a timely manner through a system that is impartial and non-discriminatory, and is free of corruption and improper government influence, all while ensuring that the rights of both victims and the accused are effectively protected.' This definition supports the proposition that both qualitative and quantitative aspects of performance should be considered when assessing the effectiveness of a judicial system.

Institutional evaluation of judicial performance

Having identified both the overarching notion (effectiveness) and the underlying aspects (independence/impartiality, competence, and efficiency) characterising judicial response to corruption, we can determine how institutional evaluation methods of judicial performance can assist us in assessing those elements.

8. Open Society Institute (2008).

Independence and impartiality

The methods used by domestic institutions to assess these interrelated features generally consist of: surveys asking citizens or experts about their perception of independence and impartiality of the judiciary; and input (also referred as *de jure*) indicators measuring the existence of specific institutions, rules, and procedures aimed at ensuring independence and impartiality.[9] The EU Justice Scoreboard developed by the European Commission provides a good example of the combined use of both categories.

Judicial independence (being one of the three dimensions assessed in the Scoreboard, together with efficiency and quality of justice systems) is assessed and measured in terms of perceived independence of courts among the general public; structural independence (for example, the existence of safeguards against the transfer of judges without their consent); and the work of the judiciary's self-regulatory bodies.

European Commission – EU Justice Scoreboard

Efficiency of justice systems	Quality of justice systems	Judicial independence
1. Length of proceedings (disposition time) 2. Clearance rate 3. Number of pending cases	1. Accessibility of justice for citizens and businesses: • Giving information about the justice system • Providing legal aid • Submitting a claim online • Communication between courts and lawyers • Communicating with the media • Accessing judgments • Accessing alternative dispute resolution methods 2. Adequate material and human resources: • Financial resources • Human resources • Training	1. Perceived judicial independence 2. Structural independence: • Safeguards regarding the transfer of judges without their consent, dismissal of judges, allocation of incoming cases within a court, and withdrawal and recusal of judges 3. Work of the judiciary's self-regulatory bodies

9. Van Dijk and Vos (2018).

Efficiency of justice systems	Quality of justice systems	Judicial independence
	3. Putting in place assessment tools 4. Using quality standards	

While relatively easy to generate, both types of indicators have received extensive criticism with regard to their validity (i.e. the extent to which they actually measure what they purport to be measuring). Perception indicators are prone to bias, as they depend heavily on the perspectives of the surveyed groups[10] and can be strongly influenced by external factors such as media coverage.[11] These limitations are all the more evident if we consider judicial response to corruption specifically. High-level corruption trials tend to polarise public opinion along the political spectrum, enhancing the risk of bias in the use of perception-based indicators. *De jure* indicators, while being objective, focus on the rules 'on the books' and give very little insight into the actual performance (i.e. output) of the assessed institution and its ability to fulfil its role in the society.

Competence

States have used a variety of approaches to evaluate the competence of judges and prosecutors in the application of the law. In many European systems, the percentage of decisions reversed in appeal and of indictments resulting in convictions are considered important criteria to evaluate the competence of judges and prosecutors respectively.[12] These two indicators, however, have serious limitations and flaws. The former could be considered to unduly promote compliance with the legal standpoints of the higher instances, therefore undermining the principle of internal judicial independence.[13] The latter is likely to have the negative effect of pushing prosecutors to bring to trial cases that are easier to prove. In the processing of corruption this would translate into an unwanted incentive to target so-called 'low-hanging fruits' rather than grand or political corruption.

10. Ginsburg (2011).
11. Mason and Johnsøn (2013).
12. CCJE (2014).
13. CCJE (2014).

Other systems have opted for softer approaches. The International Framework for Court Excellence was developed by an International Consortium consisting of groups and organisations from Europe, Asia, Australia, and the United States. It comprises a set of 'values, concepts, and tools by which courts worldwide can voluntarily assess and improve the quality of justice and court administration they deliver.' Among the values against which courts are assessed are equality before the law, fairness, impartiality, independence, and competence.

The first and main step of the process is a self-assessment which is carried out by the courts through a questionnaire aimed at evaluating seven areas of court excellence: Court Leadership and Management, Court Planning and Policies, Court Resources (Human, Material, and Financial), Court Proceedings and Processes, Client Needs and Satisfaction, Affordable and Accessible Court Services, and Public Trust and Confidence. It should be noted that this tool was not developed to assess the work of the judiciary in corruption cases. Also, being based on self-evaluation, it does not appear suitable for assessing in an objective manner a complex and politically sensitive matter as the processing of corruption cases.

Finland undertook a different and innovative method to evaluate the quality of adjudication by courts. It developed a set of 40 quality criteria aimed at assessing six aspects of adjudication: the process; the decision; treatment of the parties and the public; promptness of the proceedings; competence and professional skills of the judge; and organisation and management of adjudication. The criteria developed in relation to the second aspect include an evaluation of the competence of judges through the review of the quality of reasoning rendered in a sample of decisions. This is realised through self-evaluation by the judges; extensive survey among attorneys, prosecutors, and parties; and an evaluation by an expert group comprising a judge, an attorney, a prosecutor, a law professor, and a PR and communications professional. Though still an isolated approach, it has attracted the attention of experts and scholars developing methods aimed at measuring the quality of judicial decisions.[14]

14. Bencze and Ng(2018).

Efficiency

Compared with the previous elements, the assessment of efficiency in a judicial system is rather standardised across countries. The metrics developed in the EU Justice Scoreboard for that purpose represent a good picture of commonly accepted standards.[15] The Scoreboard adopts the following indicators to measure efficiency: length of proceedings (disposition time), clearance rate, and number of pending cases.

While these indicators are very useful to assess the overall performance of a judicial system, they do not seem sufficiently refined for the specific purpose considered here. Corruption encompasses a wide scope of conducts of different magnitude and complexity. It includes isolated episodes of petty nature as well as multi-layered schemes involving different persons in different roles.

For this reason, statistics on the number of corruption cases prosecuted and adjudicated, disregarding their specific 'weight,' would not give us a reliable picture of the level of productivity of the justice system and its overall impact in addressing corruption. Length of proceedings is similarly affected by the complexity of the case under trial. The simple average of time taken to solve a corruption case would not provide us with reliable information as to whether the judicial actors are efficiently making use of their time.

From this overview of methods and criteria to evaluate judicial systems, it is possible to conclude that they are unsuitable for carrying out a meaningful assessment of the effectiveness of judicial processing of corruption cases for two main reasons. Their goal is to evaluate judicial performance as a whole, hence they are not adapted to the specific features and challenges of corruption; and, being designed for assessing vast amounts of data, they necessarily concentrate on quantitative methods of research. Even when indicators aimed at measuring qualitative features are employed, the results are presented almost exclusively in numerical terms and offer limited insight into the institutional, normative, or behavioural factors determining a certain value or result.[16]

15. CEPEJ (2018).
16. Arndt and Oman (2006).

Notwithstanding these limitations, the data compiled by national institutions in the field of performance evaluation constitute an unavoidable starting point for any effort aimed at assessing and measuring qualitative and quantitative aspects of judicial effectiveness in a sound and comprehensive manner.

Trial monitoring

Trial monitoring is based on different premises from performance evaluation methods. It is an activity carried out by international and non-governmental organisations with a view to supporting the development of national judicial systems, often as part of broader rule of law or human rights initiatives and operations. It does so through direct observation of the functioning of critical parts of the judicial process to assess the compliance of the system with certain accepted standards.

The activities carried out as part of a trial monitoring programme may vary depending on the context, but include as a minimum the observation of public court hearings. Also, trial monitoring often relies on the review of court documents and interviews with legal practitioners to get a deeper understanding of how monitored cases are processed and the general functioning of the justice system. The data obtained through these sources are then analysed on the basis of selected criteria and the results are presented in public or non-public reports, which include recommendations to address the identified problems.

Trial monitoring is typically carried out by independent organisations and not domestic authorities. This is an important feature of this tool, as it can be used as an effective instrument to enhance public and external scrutiny on the work of the justice system in general or in specific sensitive sectors, such as anti-corruption.

While the emphasis in trial monitoring is on qualitative fact-based analysis, the methods and criteria employed are not rigid and differ considerably between programmes. This is further complicated by the fact that the methodological tenets of a specific trial monitoring programme are not usually presented in detail, at least in public documents.

The fair trial approach

A serious attempt to define the scope, methods, and techniques of trial monitoring has been realised by the OSCE Office for Democratic Institutions and Human Rights (ODIHR). The main feature of the methodology developed by ODIHR is its almost exclusive focus on the compliance of the domestic judicial process with fair trial standards, as per the relevant provisions of the main human rights treaties (i.e. International Covenant on Civil and Political Rights (ICCPR) and European Convention on Human Rights (ECHR). When criminal justice is in question, fair trial standards generally coincide with the procedural rights of the accused to a public hearing before a competent, independent, and impartial tribunal established by law; to be presumed innocent; to be tried within a reasonable time; to equality of arms; to be informed of the criminal charges; to adequate time for the preparation of the case; to be assigned a legal counsel; to call and examine witnesses; and to a reasoned judgment.

Within the defined scope of trial monitoring, the legal digest developed by ODIHR to serve as the reference manual for those involved in trial monitoring takes into consideration only the fair trial provisions of the ICCPR and ECHR, and the related case law developed by the UN Human Rights Committee and the European Court of Human Rights (ECtHR). The legal analysts monitoring the cases are provided with a detailed framework of standards and criteria related to fair trial, against which they analyse and evaluate what they see in court, read in judicial acts, or learn in interviews with judges, prosecutors, and attorneys. This means that the general criterion used to assess the quality of a judicial system consists of its level of compliance with international norms and standards on the right to a fair trial.

The Trial Monitoring Report Georgia published by ODIHR in 2014 offers a good example of the application of this methodology. After a change of government in 2012, the national judicial system initiated criminal cases against a number of high-level officials linked to the former government on charges of corruption, organised crime, abuse of office, and torture. As the Georgian authorities found themselves under increasing international scrutiny on the grounds that the trials could be politically motivated, they accepted the implementation of a trial monitoring programme by ODIHR aimed at assessing the fair trial standards in these cases.

As a result, ODIHR selected 14 cases involving high-ranking officials and, during 2013 and 2014, deployed a rotating team of 42 international trial monitors (I was one of them) who observed and reported on 327 hearings in those cases. From the data gathered, the final public report identified and analysed numerous legislative provisions and judicial practices which were deemed problematic regarding compliance with international fair trial standards. It concluded that 'the respect of fair trial rights in the monitored cases was not fully guaranteed by the Georgian criminal justice system' and proposed a number of recommendations to the national authorities.

Trial monitoring programmes based on this methodology can be very successful in addressing issues related to fair trial, such as equality of arms and the right to adequate defence. However, since they focus on the respect of the rights of the defendants, they are not well equipped to assess the capacity of a judicial system to effectively prosecute and convict perpetrators, therefore ensuring criminal accountability and victims' redress.

The human rights approach

While the fair trial-based approach has been frequently adopted in trial monitoring programmes, it is not the only model which has been theorised. The other major methodological effort in this field came from the UN's Office of the High Commissioner for Human Rights (OHCHR) as part of its series of publications dedicated to rule-of-law tools for post-conflict states and aimed at presenting a 'comprehensive overview of the principles, techniques and approaches involved in legal systems monitoring.' Legal system monitoring is broader than trial monitoring as it covers not only the work of courts and prosecutors' offices, but also law enforcement agencies, detention centres, and judicial councils.

Aside from the sectoral scope, the most relevant difference from the previous approach is in its methodological character as it relates to the broader set of principles and criteria adopted by the UN agency to carry out the monitoring. The overall approach is defined as human rights-based and refers to a number of principles which include 'non-discrimination and equal treatment, such as access to justice and fair treatment of victims, access to judicial remedy and/or redress established by law and treatment according to the domestic law, substantive and procedural fairness of

proceedings, no impunity for crimes under international law, and the independent and impartial administration of justice.'

This clearly represents a more comprehensive and ambitious approach to assessing the work of the justice system, as it makes direct reference to the need to avoid impunity and to the rights of victims along with those of the accused. Consistent with this view, the manual underlines that cases should be monitored not only from a procedural point of view but also from a substantial one, with the aim of investigating failures in prosecution, judicial bias, and the effectiveness of the entire system.

This approach could, in principle, be more fertile for the purpose of assessing judicial response to corruption. The OHCHR manual, however, differs from the ODIHR one as it does not present any sets of criteria, indicators, or benchmarks to be used to implement a human rights-based monitoring of the justice system. It is limited to listing the sources where relevant international standards can be found (namely international treaty standards; regional treaty and non-treaty standards; international customary law; and international non-treaty standards).

The human rights approach to trial monitoring has been employed with some concrete results. The Mission has been carrying out a trial monitoring programme since 2003, concentrating on different aspects of the criminal justice system. Considering the troubled wartime legacy and its influence on the post-conflict transition, one of their tasks is to monitor war crimes trials before domestic courts. In this context, it is clear that a proper assessment of judicial response to war crimes could not be limited to fair trial considerations. This is because the quest for accountability and justice for victims in these cases is as poignant as the need to ensure that defendants are not unfairly prosecuted on the basis of their ethnicity rather than their criminal responsibility. Subsequently the OSCE trial monitoring programme in BiH had to adopt a flexible approach in terms of identifying suitable criteria to analyse the data gathered through trial monitoring.

The report based on the monitoring of conflict-related sexual violence cases before the Court of BiH represents a good example of this. It is already clear from the title ('Combating Impunity for Conflict-Related Sexual Violence in Bosnia and Herzegovina: Progress and Challenges') that the emphasis is on the effectiveness of the judicial process in ensuring accountability. The main part of the report consists of its 'qualitative' assessment of the capacity of

judges and prosecutors in the interpretation of the material criminal provisions applicable in these cases. The report affirms that 'in practice, judges and prosecutors have demonstrated a sound understanding of the elements of rape, sexual slavery and other forms of sexual violence as a war crime, crime against humanity or genocide.' How did the Mission come to this important conclusion on such a significant and substantive legal point? The answer is in the report's constant reference to the extensive case law on international criminal law (including conflict-related sexual violence) developed by the UN international criminal tribunals for the former Yugoslavia and Rwanda since the mid-nineties.

These tribunals examined and interpreted each element of the above-mentioned international crimes and had a pivotal role in developing and expanding the scope of protection under international humanitarian law. Constantly referring to the standards set by those tribunals, the report considers the level of compliance of domestic legislation and judicial practice with those standards as the main criterion for assessing the ability of the national judiciary in processing those crimes.

This approach is the equivalent of using international fair trial standards as the benchmark for assessing the level of respect of the rights of the accused. This method of assessing accountability for war crimes is not isolated. The OSCE Mission to Serbia, in its trial monitoring report of 2015, similarly relied on the jurisprudence of the International Criminal Tribunal for the former Yugoslavia to assess whether the domestic courts were properly applying key legal concepts, such as superior responsibility or the connection between crime and armed conflict in their war crimes trials.

Although successful in relation to war crimes trials, this approach has inherent limitations. This is because it requires the existence of a sufficiently elaborated, consolidated, and authoritative international case law to provide the researcher with the necessary set of criteria to be applied. While this type of legal framework is available for war crimes, it is simply does not exist for other serious types of crime, such as corruption. Since the creation of an international anti-corruption court looks like an unlikely possibility, those who monitor and assess the effectiveness of judicial response to corruption (or any other criminal phenomenon which has not been the subject matter of an international court) will have to find a different answer to the quest for suitable standards.

A broader, methodological problem with the current status of trial monitoring attempts is the selective approach to the dimensions of justice which are considered in each programme or report. Sometimes, as in the Georgian programme, the attention is exclusively on fair trial; other times, as in the examples from BiH and Serbia, the analysis focusses on accountability and the rights of victims. There is little doubt, however, that a proper assessment of the effectiveness of a system would have to include both aspects, as well as issues of efficiency and productivity.

My view is that to successfully assess judicial response to corruption through trial monitoring, a new approach is needed. This would broaden the traditional scope and goals of trial monitoring by integrating it with the methods used by domestic institutions to assess the performance of their justice systems, and refining the current trial monitoring methodology by providing it with a stronger foundation in social research methods.

Conversely, the examples presented in this section demonstrate the great potential that trial monitoring activities have to address the limitations of judicial performance evaluation methods. Such activities focus on qualitative, in-depth analysis based on direct observation of the phenomena under scrutiny, as opposed to assessment based on statistical values and/or perception indicators.

The OSCE ARC project in BiH

Before presenting the methodology employed in the trial monitoring of corruption cases in BiH, it is necessary to briefly describe the domestic context. Corruption is generally recognised as endemic in BiH and a key obstacle to its development.[17] Efforts to contrast this phenomenon started in the early 2000s, but have yet to yield definitive results.[18] There are many reasons behind the failure to address this problem, but it is clear that poor judicial response to corruption is a major one. This became evident in the period between 2008 and 2013, when a number of criminal proceedings initiated against high-ranking politicians on corruption charges resulted in a series of acquittals, reinforcing the sense of impunity for grand corruption in the country.[19]

17. European Commission (2019).
18. Perry (2015).
19. Perry (2015).

The international community has been a long-term political, technical, and economic supporter of judicial reform in BiH. It was clearly interested in identifying the factors determining the outcome of these proceedings, with a view to better ascertaining the merit and possible impact of future interventions in support of domestic institutions in the rule of law and anti-corruption sector. However, no analytical fact-based assessment of the processing of these cases existed at that time. Consequently, there was no agreement as to whether the acquittals were the product of lack of investigative resources, incompetence of the judicial actors, political influence, or a combination of them.

For this reason, in 2016 the Mission, with the support of the US Department of State's Bureau of International Narcotics and Law Enforcement Affairs, decided to initiate a trial monitoring project (ARC) specifically dedicated to the issue of corruption which will end in mid-2020. Since the OSCE had been carrying out a trial monitoring programme (focussing mainly on war crimes) for many years, it was well placed for this endeavour. The challenges, however, were evident from the outset. As I was tasked to devise the methodology and the operational framework for the project, I realised very quickly that I could rely on previous trial monitoring models and methods to a limited extent only.

As explained in the previous section, the mainstream fair-trial approach was inadequate as it could not properly address issues related to effectiveness of the system in connection with accountability. The method used for the analysis of war crimes trials was not replicable due to the absence of a sovra-national jurisprudence on the criminal elements of corruption-related offences. Also, the political sensitiveness of high-level corruption cases made it very likely that any critical findings regarding their handling by the judiciary would be challenged and contested by the national authorities. The creation of an original and rigorous methodology was necessary not only to properly address the matters at stake, but also as a shield against potential controversies with the domestic powers.

The rationale of the methodology I devised to tackle these problems ('the ARC methodology') was based on two propositions: it needed to be suitable for assessing all the dimensions of judicial performance which are essential for determining the effectiveness of judicial response to corruption; and it had to employ both qualitative and quantitative analytical methods to assess these dimensions. This second aspect was important since qualitative

analysis is necessary for ensuring proper identification of the problems affecting the processing of corruption cases and their root causes. Whilst quantitative methods are required to assess broader systemic trends and to measure progress towards certain goals through time.

In the following two sub-sections I will explain how the ARC methodology realises these two propositions. Furthermore, by referring to the trial monitoring findings and recommendations presented in the two public project reports, this section will link the theoretical foundations of the project to its concrete results.

The dimensions of judicial response to corruption: Identification and analysis

In the section on 'Defining effective judicial response to corruption,' I recognised competence, independence, and impartiality, as well as productivity and timeliness, as fundamental features determining the effectiveness of the judiciary in the processing of corruption cases. To assess these matters through trial monitoring, the ARC methodology identifies four dimensions (productivity, competence, fairness, and efficiency) of judicial response to corruption, and a set of criteria and guidelines for analysing relevant data obtained through trial monitoring in connection with each of these dimensions.

Productivity

This element is usually considered in the realm of efficiency. The decision to give it a distinct space was based on the fact that, in the ARC methodology, productivity takes into consideration not only the quantity but also certain qualitative aspects of the cases. The resources and time necessary to investigate and try high-level corruption are not comparable with those required for petty corruption.

As a result, productivity should be assessed in terms of corruption cases initiated and completed by the justice system in a given year, according not only to their quantity but also their complexity and seriousness. This would confirm the extent to which the prosecution is using its resources to tackle serious and systematic forms of corruption as opposed to pursuing charges in petty corruption cases. Under this notion of productivity, a decrease in the number of cases processed would not necessarily indicate a negative trend if

coinciding with an appropriate increase in the level of seriousness of those cases.

To assess these aspects consistently and reliably, cases monitored under the ARC project are categorised as high, medium, or low level on the basis of two criteria: status of the accused and gravity of the (alleged) conduct. Based on detailed instructions, a score from one to three is assigned for each of the criteria and the final score determines whether the case is high, medium, or low level according to its overall level of seriousness. In the initial stages of the project, the methodology considered a two-category system (mirroring grand and petty corruption).

After testing, however, it became clear that this categorisation would not be sufficiently refined to capture the value of a relevant number of cases which, although not classifiable as grand corruption, are nevertheless substantially more complex and/or politically sensitive than petty corruption cases. Therefore, I eventually opted for the current three-category classification. The categorisation exercise is crucial not only for the assessment of productivity but also in relation to a key requirement for any trial monitoring project, i.e. the adoption of clear and transparent criteria for the selection of cases to be monitored.

Trial monitoring, involving the observation of trial hearings, is resource consuming. As in most countries, the monitoring of every corruption case in BiH would not be feasible. For this reason, a categorisation system is essential to select the cases to be monitored in a consistent way. In the ARC project, the OSCE decided to focus on serious cases of corruption. As a result, it undertook to monitor all high-level and medium-level cases tried in the country, while monitoring only a selection of petty corruption cases, to the extent allowed by its resources.

The application of this system enabled the trial monitoring project to reach important conclusions regarding the status of judicial response to corruption. In BiH, corruption cases are tried not only before courts of general jurisdiction, but also before courts which, in principle, should have special jurisdiction on more serious forms of corruption, economic crimes, and organised crime. This institutional design, however, proved to be in conflict with reality, as trial monitoring findings showed that the majority of high- and medium-level cases processed in the country were not tried before the specialised courts but before other courts having general jurisdiction.

This observation pointed to another problematic finding: that the application of the jurisdictional criteria for the allocation of cases in BiH is flawed.

Identifying this problem and its causes, allowed the OSCE to propose some fact-based recommendations with regard to the desirable division of roles between ordinary and specialised courts in the processing of corruption. These findings also influenced the decision of some donors in identifying which courts and prosecutors' offices could mostly benefit from projects in the field of support to the domestic judiciary in processing corruption cases.

Competence

Reconsidering the concept of productivity in connection with judicial response to corruption required some thinking, but was not as challenging as assessing the level of competence of judges and prosecutors in the application of the law in the monitored cases. The difficulty was mainly because of the highly qualitative nature of this type of assessment, which made it more at risk of subjectivity and bias by those carrying out the monitoring.

The ARC methodology tackles this problem by limiting the focus of the assessment to the quality of the most important judicial acts issued by judges and prosecutors in their respective capacity: first instance verdicts and appeal decisions by the former, and indictments by the latter. This approach is justified on the basis of theoretical considerations which I present below.

Many experts consider the quality of written verdicts and appeal decisions to be the key aspect of quality of justice in general,[20] as respect for the rule of law requires that criminal laws are interpreted in a uniform and predictable way to ensure accountability, legal certainty, and equality. This is best demonstrated through judicial decisions and their reasoning in particular. A coherent and convincing reasoning is a strong guarantee against arbitrariness. A flawed, unclear, or unconvincing reasoning may indicate a lack of competence or the presence of illegitimate or non-legal considerations in the mind-set of the judge. So, a well-reasoned judgment can be taken as a reliable indicator not only of the competence but also of the independence of judges.

20. Bencze and Ng Gar (2018).

While it is easy to agree on the importance of proper reasoning for the quality of justice, it is much more difficult to define the correct criteria and boundaries for assessing this element. In particular, the evaluation of the quality of judicial decisions can, and should be, kept distinct and separate from any process aimed at checking and second-guessing the reconstruction of facts and interpretation of the law by judges. This is a scope of review which can be properly conducted by an appeal court. However, its adoption in the context of trial monitoring would overstep the limits of appropriate oversight over the work of the judiciary and arguably impinge upon its independence.[21]

Avoiding this risk requires identifying and applying relevant, clear, and authoritative standards which characterise a convincing legal reasoning. In this regard, the opinion of the Consultative Council of European Judges (CCJE) can offer valuable guidance. This body of authoritative experts holds the view that an evaluation of the competence of judges through the review of the quality of reasoning is possible without impinging on their independence.

An evaluation of the competence of judges through the review of the quality of reasoning is possible without impinging on their independence.

It also illustrates the main criteria to be considered for that purpose. The reasons must be consistent, clear, unambiguous, and not contradictory; must allow the reader to follow the chain of reasoning which led the judge to the decision; must respond to the parties' submissions (i.e. to their different heads of claim and to their grounds of defence); and should refer to the relevant provisions of the Constitution or relevant national, European, and international law, and where appropriate, to national, European, or international case law.

The opinion of the CCJE and the Finnish experience in evaluating the competence of judges through the review of the quality of reasoning,

21. Bencze and Ng Gar (2018).

convinced me that these criteria could be relied on to assess the competence of judges as part of the trial monitoring of corruption cases.

The ARC methodology adopts a mirroring approach with regard to prosecutors, as it takes the quality of indictments as meaningful and reliable evidence of their competence and develops a set of criteria to ensure reliability and consistency in the assessment. The core criterion refers to the notion of 'charging accuracy.' According to the European Court of Human Rights, it can be confirmed that a prosecutor has been accurate in the filing of charges when the charges clearly and comprehensively illustrate 'the cause of the accusation, that is to say, the acts he is alleged to have committed and on which the accusation is based, but also of the nature of the accusation, that is, the legal characterisation given to those acts.'

The first ARC public report contains a detailed analysis of the level of competence of judges and prosecutors, mainly (but not exclusively) through assessment of the quality of the main legal acts issued by them in the monitored cases. Its conclusions on the matter are critical.

With regard to the quality of prosecution, it found that 'in a number of cases analysed for the present report, the description of the criminal behaviour in supporting the charges in the indictment was seriously flawed due to the lack of, or unclear identification of, one or more of the elements of the offence.' Importantly, the report noted the strong correlation between poor quality of indictment and subsequent acquittal in high-level cases of corruption as in medium- and low-level ones.

There is strong correlation between poor quality of indictment and subsequent acquittal in high-level cases of corruption, as in medium- and low-level ones.

The criticism, however, was not limited to the work of prosecutors. The report also established that in the monitored cases, judicial decisions were often based upon unclear or insufficient reasoning. Specifically, 'in some cases, flaws in the reasoning were related to the manner of presentation and evaluation of the evidence as it was not assessed in light of the elements of the crime. In other cases, the reasoning was not structured in a way that the

elements of the crime were identified and addressed separately.' It also noted how judges, in their reasoning, seldom referred to jurisprudence to justify their stances on the interpretation of contested norms. This is one of the main factors behind the inconsistent application of key criminal provisions from one case to another.

The identification of these problems was accompanied by the proposal of measures and recommendations aimed at improving the competence of judges and prosecutors in the application of the law. For example, the report recognised that the poor use of jurisprudence in reasoning was due not only to cultural and educational reasons, but also the lack of efficient search tools. Subsequently, it recommended the creation of a database for researching jurisprudence by topic.

Fairness

The third dimension of judicial response to corruption refers to the fairness of the process, particularly in terms of adherence to fair trial standards. Respect of these standards is not usually assessed separately in the context of judicial performance evaluation. Instead, the correlated aspects of independence and impartiality are taken into consideration (typically through perception surveys). Therefore, the decision to give prominence to fairness in the ARC methodology was justified by the evident importance of this factor for the proper administration of justice in general and, particularly, in the field of judicial response to corruption. The effectiveness of a judicial system cannot be evaluated only in terms of its ability to punish criminals; acquitting the innocents is an equally important feature. The level of respect of fair trial rights can be considered as the best evidence of this side of the effectiveness equation.

Fair trial standards include what is arguably the best set of criteria for assessing in a direct, fact-based manner the level of independence and impartiality of the judiciary. Specifically, they provide a framework to look for signs of bias in the actual behaviour of a judge in court or the existence of objective facts (eg conflict of interests) which raise doubts as to their impartiality. Also, it can be argued that respect of the rights of the accused is a strong sign of the impartiality of the judiciary, especially when individuals who are not connected to the government are under trial. These considerations are all the more valid in connection with the prosecution of corruption, as this can be used as an instrument for persecuting political

opponents in systems where the judicial system is influenced by the forces in power.

Assessment of this aspect did not present the previously noted methodological dilemmas. As seen in the section on 'Trial monitoring', fair trial standards are at the core of mainstream trial monitoring approaches. Accordingly, the ARC methodology relies on the criteria and methods developed in that context as relevant and adequate for the monitoring of corruption cases, i.e. while the assessment of competence is mainly based on the review of written acts, the analysis of fairness draws most of its data from the direct observation of hearings. Since this is where the respect of the rights of the accused (publicity of trial, equality of arms, right to defence, etc.) can be best observed.

Through trial observation, the OSCE was in the position to determine that the behaviour of judges in court was, with few exceptions, free of visible bias. However, it ascertained certain problems in applying provisions aimed at regulating the grounds for disqualification of a judge when doubts are raised regarding the impartiality of an entire court rather than in relation to a single judge. Trial monitoring analysis, informed by reliable data and clear international standards, identified the origins of this problem in a legislative gap, and recommended the amendment of the relevant provisions to solve it.

Efficiency

Since efficiency in terms of productivity is considered in the first dimension above, in the ARC methodology this notion refers exclusively to timeliness of the proceedings. This factor is prominent in every institutional evaluation of judicial performance and is assessed through the adoption of indicators, such as disposition time, and the setting of benchmarks, such as optimal and predictable timeframes for the processing of cases.

These methods are suitable for identifying the phases of the process and the types of cases where delays are more frequent. Trial monitoring, though, allows for a deeper look at the factors and behaviours causing these delays, for example the specific reasons for the adjournment of hearings or the skills of the judges in organising the trial from a managerial point of view.

The analysis of trial monitoring data led to the conclusion that 'primary factors affecting the length of proceedings in corruption cases occur at the trial stage and concern 1) changes in the composition of the panel requiring

a recommencement of the trial; and 2) poor management of the trial by the presiding judge, especially with regard to the use of available measures to ensure the presence of parties at the trial.' It also pointed to the fact that delays and related management problems were more pronounced in high-level corruption cases. As a result, the report recommends the adoption of guidelines for the management of trials and plea hearings in complex cases.

Quantifying the effectiveness of judicial response to corruption: The search for indicators

In the previous sub-section I presented how the ARC methodology identified and assessed the main factors determining the effectiveness of judicial response to corruption. This required the elaboration of an innovative trial-monitoring methodology (enriched by referencing and reshaping of concepts such as quality of justice, competence, productivity, and timeliness which have been explored in a different context): institutional evaluation of judicial performance.

The ARC methodology also draws from the practices developed in that context in relation to the significance given to quantitative research methods. It is important to note that trial monitoring programmes traditionally privilege qualitative analysis based on a mainly legal conceptual framework. Although statistics and other numerical data are often used, this is not done in a consistent manner, but instead to support or provide context to the qualitative findings.

My conviction, however, is that a comprehensive trial monitoring methodology should not only evaluate but also measure the effectiveness of judicial response to corruption. Qualitative analysis should provide a detailed picture of the main factors and features contributing to a certain situation or performance during a certain period of time. Although, this approach is not suitable for measuring progress (or its absence) towards certain goals or the impact of reforms undertaken through time. Quantitative analysis is necessary for this purpose as it can be replicated in a reliable manner, allowing comparison through time and producing more 'generalisable' results.

For this reason, the ARC methodology envisages the development of a set of indicators of effectiveness of judicial response to corruption cases. The

first phase of the project focussed on the qualitative identification and assessment of problems, as well as the proposition of appropriate measures, and resulted in the two public reports. The definition and testing of these indicators is currently ongoing as it constitutes the main activity of the last phase of the project, which should be concluded in April 2020.

While this is still work in progress, the main features of the indicators are already well established. It will measure both quantitative and qualitative aspects of judicial response to corruption. Each of the four dimensions will have its own set of indicators. The indicators will be based on both trial monitoring data and official statistics in a balanced and complementary manner. Recognising the inherent deficiency of even the most sophisticated indicator taken in isolation, taken as a whole and in their interconnection, they aim at reducing the ambiguity of each indicator by balancing it against the others.

For example, the ambiguity of institutional indicators of productivity based on number of initiated cases will be reduced by indicators measuring the level of seriousness and complexity of these cases. The ambiguity of competence indicators based on the percentage of acquittals or quashed verdicts will be reduced by indicators measuring the quality of indictments, verdicts, and appeal decisions. Finally, the indicators will have a high degree of reliability as they will be based either on objectively quantifiable data (for example, number of indictments and verdicts) or on qualitative information where the correlated risk of subjectivity will be limited by their being based on direct observation of hearings and review of documents rather than perceptions.

Impact and limitations of the ARC project

Although in its final stage, it is possible to give an evaluation of the project's impact and limitations. The first public ARC report published in February 2018 was based on an analysis of 67 corruption cases completed by the domestic judicial system and monitored by the OSCE in the period between January 2010 and September 2017. It contained 15 recommendations to tackle different problems identified through trial monitoring applying the ARC methodology.[22] The second report, published

22. Since the cases had been monitored before the start of the ARC project, the data from the observation of hearings were not gathered according to the ARC methodology. However, the

in April 2019, was based on a broader sample of 300 cases monitored 'live' in the period between 2017 and 2018 and included nine additional recommendations.

These recommendations were addressed to different institutions, including the executive and legislative powers, courts and prosecutors' offices, and the high judicial and prosecutorial council. Although the report's findings were critical of the institutions in charge of ensuring an effective processing of corruption cases, the recommendations were generally accepted by the domestic authorities as relevant and objective.[23]

The authorities acknowledged the problems identified by an independent observer in a sensitive area such as the processing of corruption cases, and this can be taken as an indication of the reliability and soundness of the methodology. Different precautions had been taken to ensure the accuracy of the findings and minimise the risks of bias and subjectivity in the analysis of data deriving from the observations of hearings and the review of judicial acts. These precautions included multiple peer-review processes as well as the engagement of external experts to confirm or challenge the findings.

Acceptance by the national counterparts should not be taken *per se* as evidence of a genuine commitment of the authorities to improve the effectiveness of judicial response to corruption. In this regard, the second ARC public report published in 2019 emphasised that no sign of progress could be detected in addressing six out of the 15 recommendations. The reduced impact of the recommendations (at least in the short term) points to an intrinsic limitation of the project – no matter how targeted and feasible the proposed measures, their implementation remains mainly, if not exclusively, in the hands of the domestic authorities.

However, the project does not involve the monitoring of the proceedings in their entirety. Under the ARC methodology, corruption cases are monitored from the moment the procedure is public, either at the time filing the

---

new methodology was applied to the other key component of trial monitoring, i.e. the review of the quality of judicial acts.

23. As recognised in the second ARC report, 'the vast majority of the report's recommendations were endorsed by the domestic authorities in the form of several conclusions at a roundtable organised by the European Commission in June 2018.' Additionally, 'in September 2018, the HJPC endorsed the roundtable conclusions addressed to it by embedding them in a specific Action Plan on the Fight against Corruption and Organized Crime, including Money Laundering.'

indictment, or at the time of the suspect's arrest when pretrial detention is sought. This means that it does not observe directly the investigative stage of the process. This is an important limitation of the project as many corruption cases – including high-level ones – never reach the indictment phase as they are closed for lack of evidence after the investigation.

Many corruption cases – including high-level ones – never reach the indictment phase as they are closed for lack of evidence after the investigation.

This restriction is common to many trial monitoring projects and is due to the difficulty and sensitiveness of obtaining access to official information in the investigation stage because of its confidential nature. The ARC project tackled this problem by assessing the quality of investigations indirectly through an analysis of the quality of indictments, including the evidence presented by the prosecution. This approach resulted in some useful insight on the link between quality of the investigation and outcome of the case in court. It is, however, limited as it does not enable an examination of cases which do not reach the indictment due to lack of evidence.

Another limitation of the project is that it does not consider the impact of external factors on the processing of corruption cases, such as the type of media reporting on the trials, or pressures on and threats to members of the judiciary. Environmental factors can be very important as the political and social context plays a pivotal role in processing both high-level and petty corruption cases.

Measuring judicial response to corruption across countries

Improving the effectiveness of judicial response to corruption depends on many factors, including the political will of the national authorities and their capacity to reform dysfunctional institutions.

While trial monitoring programmes cannot produce positive changes directly, they have great potential to advance how we evaluate and measure the effectiveness of judicial response to corruption. The importance of this endeavour cannot be underestimated. It is important to note that the European Commission's 2018 Strategy for the Western Balkans stresses that, in order to accede to the EU, this region will have to take effective steps to counter corruption and demonstrate 'concrete and sustained track record in tackling corruption, money laundering and organised crime'.

Trial monitoring programmes have great potential to advance how we evaluate and measure the effectiveness of judicial response to corruption.

On the EU side, the Strategy calls for 'more detailed rule of law assessments' and more systematic 'monitoring of implementation and enforcement.' It introduces trial monitoring as one of the tools to be employed for the purpose of evaluating progress in judicial reform and effectiveness of the justice system. This is remarkable considering that trial monitoring is new to their enlargement process toolbox. These developments are not limited to policies towards States in the accession phase. In a 2019 Communication, the European Commission presented its plan to 'deepen its monitoring of rule of law related developments in the Member States,' including the ability of these States to counter corruption.

Therefore, as the potential of trial monitoring has been recognised at the policy level, this paper aimed to give a concrete example of how these types of activities can be implemented in practice to assess judicial response to corruption. The methodology developed for the OSCE ARC project in BiH represents a new approach to trial monitoring, as it blends methods and concepts deriving from that field with those used in the context of institutional evaluation of judicial performance. Its implementation has been successful so far, as it has provided national and international policymakers with a fact-based and comprehensive assessment of the problems and factors determining judicial response to corruption.

It is argued that this approach to trial monitoring can be adopted for similar purposes in other national contexts and can be implemented by international

or civil society organisations active in the field of anti-corruption and judicial reforms. As I attempted to demonstrate, the four dimensions of judicial response to corruption identified in the ARC methodology, as well as the correlated criteria devised for their assessment, are relevant and applicable in each jurisdiction. However, it must be acknowledged that its implementation in other jurisdictions could be hampered by limitations in the access to judicial acts, such as verdicts and indictments, and to court hearings, for example owing to the extensive use of closed sessions.

Moreover, trial monitoring is resource consuming as it requires direct observation of hearings by individuals with legal skills. In this regard, the categorisation criteria of the ARC methodology can be used to limit the focus of the monitoring to high-level corruption cases. Such focus would make the implementation of a trial monitoring project manageable from a resource point of view, especially in those countries where the trial of high-level cases takes place before a specialised court.

Based on the experience in BiH, it is reasonable to expect that a team of five persons, properly trained, can cover trials in 30–40 cases in one year, which can be considered a sufficient sample to draw some general conclusions. Given that cases may take an additional year or more to be finalised in the appeal stage after the public trial, trial monitoring projects should be planned to last for a minimum of three years to allow for an assessment of the entire course of the selected proceedings. The qualitative data derived from the monitoring of a limited number of trials could be complemented with findings from the analysis of official statistics. The ARC methodology demonstrates how these two types of sources can be integrated to provide a fuller picture. Also, the use of indicators makes it appealing in the context of projects aimed at multi-country or regional assessments of judicial response to corruption, as it would enable the comparison of results in the different jurisdictions.

References

ADB-OECD Anti-Corruption Initiative for Asia-Pacific Combating Corruption in the New Millennium. 2003. Effective prosecution of corruption.

Albers, P. The assessment of court quality: Hype or global trend?

Arndt, C. and Oman, C. 2006. Uses and abuses of governance indicators. Development Centre. OECD.

Bencze, M. and Ng Gar, Y. 2018. How to measure the quality of judicial reasoning. Springer.

CCJE. 2008. Opinion no. 11 on the quality of judicial decisions.

CCJE. 2014. Opinion no. 17 on the evaluation of judges' work, the quality of justice and respect for judicial independence.

CEPEJ. 2018. European judicial systems: Efficiency and quality of justice.

ECtHR. European Convention on Human Rights.

ECtHR. 1999. Pélissier and Sassi v. France.

European Commission. EU Justice Scoreboard.

European Commission. 2018. A credible enlargement perspective for and enhanced EU engagement with the Western Balkans.

European Commission. 2019. Commission opinion on Bosnia and Herzegovina's application for membership of the European Union.

European Commission. 2019. Strengthening the rule of law within the Union: A blueprint for action.

Ferguson, G. 2018. Global corruption: Law, theory and practice. University of Victoria.

Ginsburg, T. 2011. Pitfalls of measuring the rule of law. Hague Journal of the Rule of Law 3(2).

Mason, P. and Johnsøn, J. 2013. The Proxy Challenge: Why bespoke proxy indicators can help solve the anti-corruption measurement problem. U4 Brief 2013:2. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

Mohr R. and Contini F. 2007. Judicial evaluation in context: Principles, practices and promise in nine European countries

National Centre for State Courts. 2013. The international framework for court excellence.

ODIHR. 2012. Trial monitoring: A reference manual for practitioners.

ODIHR. 2012. Legal digest of international fair trial rights.

ODIHR. 2014. Trial monitoring report Georgia.

OHCR. International Covenant on Civil and Political Rights.

OHCHR. 2006. Rule-of-law tools for post-conflict states: Monitoring legal systems.

Open Society Institute. 2008. Promoting prosecutorial accountability, independence and effectiveness.

OSCE Mission to BiH. 2014. Combating impunity for conflict-related sexual violence in Bosnia and Herzegovina: Progress and challenges.

OSCE Mission to BiH. 2018. Assessing needs of judicial response to corruption through monitoring of criminal cases (the ARC project).

OSCE Mission to BiH. 2018. Assessing needs of judicial response to corruption through monitoring of criminal cases (ARC) – Project report.

OSCE Mission to BiH. 2019. Assessing needs of judicial response to corruption through monitoring of criminal cases (ARC) – Trial monitoring of corruption cases in Bosnia and Herzegovina: Second assessment.

OSCE Mission to Serbia. 2015. War crimes proceedings in Serbia (2003–2014).

Perry, V. 2015. A cross-cutting survey of corruption and anticorruption issues in Bosnia and Herzegovina: Overview, challenges and recommendations. USAID.

Rovaniemi Court of Appeal. 2006. Quality project of the courts in the jurisdiction of the Court of Appeal of Rovaniemi, Finland: Evaluation of the quality of adjudication in courts of law.

Simion, K. 2016. Qualitative and quantitative approaches to rule of law research. INPROL.

Stephenson, M. C. and Schütte, S. A. 2016. Specialized Anti-Corruption Courts: A Comparative Mapping. U4 Issue 2016:7. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

Stephenson, M. C. and Schütte, S. A. 2019. An International Anti-Corruption Court? A synopsis of the debate. U4 Brief 2019:5. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.

United Nations. International Criminal Tribunal for Rwanda.

United Nations. International Criminal Tribunal for the former Yugoslavia.

Van Dijk, F. and Vos, G. 2018. A method for assessment of the independence and accountability of the judiciary. International Journal For Court Administration, 8(2).

Vera Institute of Justice. 2003. Measuring progress toward safety and justice.

World Justice Project. Criminal justice (Factor 8).