

Putting Our Best Foot Forward: The Judiciary and the Challenge of Satisfying Justice Needs of the 21st Century

Abuja, Nigeria 7-9 July 2014

“Judicial Performance Evaluation: Perspectives from Europe”

Dr Gar Yein Ng

Independent researcher

Introduction

The paper will describe experiences of performance evaluation of judges in the European context, examining what systems are being used across the Continent; how these are undertaken; how these models compare with one another in terms of strength and effectiveness; what difference performance evaluation has made to productivity outcomes, and how these have helped to ensure fairer, speedier delivery of justice services.

The first section will outline a theoretical framework laying out what performance evaluation is, what norms are important under international law; and what the regional guidelines in Europe are on applying judicial performance evaluation. Following from this will be a comparative examination of various European countries, looking at both established democracies and transitioning democracies. Finally, brief consideration will be given to the possibilities of judicial performance evaluation in Nigeria.

“Theoretical Framework¹”

1. General performance evaluation: an introduction

The foundation for performance evaluation in public services has come as a by-product of developments within requirements to good governance, as well as new public management and information technologies. Performance evaluation of the courts as such has been one tool used to hold courts to account for public spending and access to efficient delivery of justice. Performance evaluation of individual judges on the other hand, is a relatively recent phenomenon, used for different purposes than performance evaluation of courts. In practice, performance evaluation of courts is very much dependent on objective and quantitative criteria, such as settlement rates; average waiting times for cases to be heard; costs per case; physical and virtual accessibility and so forth. These criteria take into account performance of all staff and how efficient a court has been. Quantitative criteria are therefore considered as appropriate measure for court evaluation.

However, performance evaluation of judges looks only at the individual judge, focussing on professional skills, such as quality of reasoning, comportment, compliance with ethics, and computing skills. In practice, it has developed in the last decade across Europe (Belgium, France,

¹ This section is taken from the “Assessment of the Performance Evaluation of Judges in Moldova” June 2014 OSCE/ODIHR

Netherlands, Italy, Spain, Austria and Germany- all of which have statutes allowing for evaluation). The purpose of a particular system for performance evaluation of judges affects the type of criteria used for evaluation and the way data is collected in that system.

The following describe generally the main purposes of performance evaluation of judges: to improve the performance of individual judges and enhancement of professional accountability of judges; to enhance public confidence; and to aid judicial institutions in deciding upon issues of careers and promotions within the judiciary. In some rare cases, performance evaluation is linked with disciplinary and dismissal proceedings. These purposes are not exclusive to each other, and any model may take on more than one purpose.

2. International norms on judicial independence

International standards of judicial independence require that judges' positions must be protected against outside influences, in order to ensure that judgments are fair and unbiased.

This part examines the characteristics of judicial independence in international law. Judicial independence is an institutional requirement in international law.²

International norms, such as article 6 of the European Convention on Human Rights, article 10 of the Universal Declaration of Human Rights, and article 14 of the International Covenant on Civil and Political Rights emphasise the right to an independent and impartial tribunal as part of the right to a fair trial:

“Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal established by law.”³

The principle of judicial independence has been further fleshed out by the Council of Europe in documents such as the Recommendation CM/Rec(2010)12, or the OSCE, in the Kyiv Recommendations. The purpose of this principle is to protect the right to a fair trial, be it in civil, criminal or administrative law, by protecting judges from improper influence, both external and internal to the judiciary.⁴

It is important to have this background in mind when assessing any performance evaluation system for judges that it does not interfere in the decision making of the judge (i.e. that evaluators do not seek to encourage certain case decisions for certain evaluation outcomes of judges; nor that any government or politician be able to influence the evaluators to inflate or deflate the grades of judges).

² See UN Basic Principles on the Independence of the Judiciary; for domestic examples, see Constitution of France article 66, Constitution of USA, article III; Germany's Basic Law article 97 etc. In the OSCE, according to Ministerial Council decisions of Copenhagen (1990), Moscow (1991), and Helsinki (2008) there is a political commitment to respect judicial independence as recognized in international standards such as the UN Basic Principles, the International Covenant on Civil and Political Rights, and the Universal Declaration of Human Rights. A compilation of all OSCE human dimension commitments is available at <http://www.osce.org/odihr/76902>.

³ Article 6, European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR).

⁴ Council of Europe Recommendation 2010/12. Para 22.

International standards, as reflected in the Kyiv Recommendations,⁵ and Venice Commission Opinion on Armenia’s draft law on judicial performance evaluation are examined here. There is recognition at an international level of the fact that performance evaluation of individual judges is important on the one hand to strengthen judicial independence, capacity and professionalism, and on the other hand to lend transparency and accountability to judges and their work.⁶

3. Judicial independence and Performance Evaluation of Judges

i. Aims within the European Context

There is increasing international recognition for the need to develop merit based systems of judicial promotion to create legitimacy and confidence in judges and courts, and to lend legitimacy to the procedures of promotion within courts and to higher courts.⁷

The aims of such systems may vary, as discussed above, between promotion and careers, self-improvement, public confidence and more rarely: discipline. These various aims and how they can be met while upholding judicial independence was discussed at an expert meeting in Kyiv on “Judicial Independence in Eastern European, South Caucasus and Central Asia: Challenges, Reforms and Way Forward” in June 2010 that led to adoption of the Kyiv Recommendations.⁸ The Venice Commission has also examined these aims and how they interact with the principle of judicial independence.

Concerning using the results of performance evaluation to decide on issues of careers and promotions, the Venice Commission finds that it is a controversial issue, and suggests that promotion decisions do not rely solely on the results from performance evaluation.⁹

If the primary focus of the evaluation system is skills building, it has been argued that it will more likely have the cooperation of the judges. “If it is seen as leading to consequences such as exclusion from promotion, that co-operation may not so willingly be given. Evaluation should not be seen as a tool for policing judges, but on the contrary, as a means of encouraging them to improve, which will reflect on the system as a whole.”¹⁰ Performance evaluation of judges should be clearly separated from the system of disciplinary responsibility of judges. In its opinion on the draft legislation of Armenia on their proposed evaluation system, the Venice Commission has also highlighted that “The fact that a negative result in the evaluation procedure could lead to a disciplinary sanction, including dismissal, is problematic.”¹¹ In line with

⁵ The Kyiv Recommendations on Judicial Independence in Eastern Europe, South Caucasus and Central Asia were developed by a group of independent experts at a regional expert meeting organized by ODIHR and the Max Planck Institute for Comparative Public Law and International Law – Minerva Research Group on Judicial Independence, in June 2010; available also in Romanian at <http://www.osce.org/odihr/KyivRec>.

⁶ See e.g. OSCE/ODIHR; Max Planck Institute for Comparative Public Law and International Law, 'Judicial Independence In Eastern Europe, South Caucasus And Central Asia: Challenges, Reforms and Way Forward' 2010 para 28; and United Nations Office on Drugs and Crime, 'Resource Guide on Strengthening Judicial Integrity and Capacity' 2011 p.5.

⁷ UN Basic Principles on the Independence of the Judiciary, article 13; Council of Europe Recommendation on Judicial Independence, 44-48 & 58; and European Charter on the Statute for Judges, article 4.

⁸ OSCE/ODIHR; Max Planck Institute for Comparative Public Law and International Law, 'Judicial Independence In Eastern Europe, South Caucasus And Central Asia: Challenges, Reforms and Way Forward' 2010.

⁹ Venice Commission Opinion Armenia para. 94.

¹⁰ Venice Commission Opinion Armenia para 23.

¹¹ Ibid. para 9.

international standards for judicial independence, performance evaluation should never be used to assess the content of decisions and verdicts and certainly not as a basis of sanction.¹²

ii. Criteria

The Kyiv Recommendation¹³ highlight that performance evaluation of judges should be primarily qualitative in nature, and focus on the skills required to be a judge. The following can be considered as qualitative skills: professional competence, which assesses knowledge of both procedural, substantive and evidentiary law, the ability to conduct trials, and the capacity to write reasoned decisions; personal competence, referring to a judge's ability to cope with workload, the ability to decide cases, and an openness to using new technologies in their function; and social competence, which assesses judges' ability to mediate and show respect for parties. These are not controversial and go to the heart of the judicial function.

Quantitative performance evaluation is more controversial as there is a danger that there will be a focus only on productivity, compromising quality in favour of quantity of judges' work. The Kyiv Recommendations state that this benchmark is only to be used for self-improvement for judges, and should not be focussed on as the main element of evaluation of a judge's overall performance. Furthermore, it is the opinion of the Venice Commission that "quality, within a reasonable period of time and fairness are the necessary prerequisites that the judge should meet when rendering a decision."¹⁴ It re-states the Kyiv Recommendations that counting cases and reversal rates "should not be used to the detriment of the individual judge", but rather as a tool of judicial administration.¹⁵

A number of international recommendations approve of evaluating the quality of conducting hearings and writing decisions.¹⁶ In evaluating the quality of decisions, a practice evolved of measuring the level of reversals, alongside other qualitative criteria such as reasoning, and legibility. The danger of calculating the numbers or rates of reversals is that it "... is likely to produce a timid judiciary which looks over its shoulder all the time..."¹⁷ One consequence of this is that judges could start asking judges of the higher court how to decide their cases in a way that would avoid reversals. This limits access to justice at first instance, and creates interference in decision making by the higher courts.¹⁸ A further extreme consequence of this would be sanctions against judges for the content of the decisions.¹⁹ However, sanctions for the content of judges' decisions and their interpretation of the law must be avoided or strictly limited.²⁰

Another practice that has developed over recent years, though connected more to court performance than individual judicial performance, has been the use of customer surveys. These can be very relevant to individual performance, especially for enhancing judicial comportment towards litigants. If done with respect to judicial independence, it is a useful tool to highlight

¹² Kyiv Recommendations, para 28.

¹³ Paras 27-29.

¹⁴ VC Opinion Armenia, para 12.

¹⁵ Ibid. para 34 Questioned again on procedure in para 39.

¹⁶ Kyiv Recommendations, para 31.

¹⁷ VC Opinion Armenia. para 40.

¹⁸ Ibid. paras 18 and 39.

¹⁹ Kyiv Recommendations, paras 25 and 28.

²⁰ Council of Europe Recommendation on Judicial Independence, CM 2010(12), paras 66 and 68.

issues such as communication between judges and litigants; managing expectations of litigants. It is also a useful tool to give a voice to litigants and parties as to what they need from their courts.²¹ The use of such surveys is also important to highlight where the actual problems lie in terms of judicial performance: issues of concern to various court users in Europe have included “accessibility (cost and complexity), delay, fairness and judicial competence...”²²

All statistical data, whether pertaining to reversal rates or settlement rates also needs to be read in context; for example in terms of reversal rates, there are situations which are not attributable to the judge. The Venice Commission also questions the assumption that the higher court always gets it right, highlighting that the first instance court decision quashed by the court of appeal could well be supported by a court of cassation, constitutional court or the European Court of Human Rights.²³ Also, when analysing statistics for settlement rates and duration of trials, data should be analysed in light of the complexity of cases, case loads of judges (also between courts), and sudden increases in caseloads, amongst other things.²⁴ The Venice Commission acknowledges that problems with efficiency are of concern when it comes to the right to a fair trial. Nevertheless statistics on efficiency and settlement rates should only be used as a starting point to identify any possible problems for a judge.²⁵ If a system is over-relying on such statistics, this may tempt judges to ‘disregard what would normally be seen as necessary under the law and his or her interpretation of it.’²⁶

iii. Transparency and fairness of evaluation procedures

International standards also provide recommendations on the transparency and fairness of performance evaluation systems. Recommendations on transparency relate to the need to publish the criteria, and to conduct the procedure of evaluation in a transparent manner.²⁷ In most countries that have a system of performance evaluation of judges, it is governed by publicly accessible laws and regulations, specifying the criteria, indicators and how to measure them, and procedures for reaching evaluation decisions. Such transparency makes the evaluation and possible consequences foreseeable for judges and therefore more fair and credible. Transparency is also important where evaluation is conducted to enhance public trust in the judiciary, and to counter perceptions of dependence, undue influence and corruption. Recommendations on the fairness of procedures include the judges’ rights to be heard and to appeal.²⁸

Methodology

This paper reflects a European perspective using comparative methodology. New material is recently available from the Council of Europe’s Consultative Council for Judges in Europe, which is soon to publish an opinion about performance evaluation of judges on the basis of

²¹ Contini, Francesco and Mohr, Richard, *Judicial evaluation : traditions, innovations and proposals for measuring the quality of court performance*, VDM Verlag Dr. Müller, Saarbrücken 2008 p.17; and pp. 72-74

²² Ibid. p.72

²³ VC Opinion Armenia, paras 39-40.

²⁴ Ibid. para 37.

²⁵ Ibid. para 37.

²⁶ Ibid. para 43.

²⁷ Kyiv Recommendations, paras 29 and 31; VC Opinion Armenia, para. 70.

²⁸ Kyiv Recommendations para 31. VC Opinion Armenia, para 86.

responses to questionnaires sent to member states of the Council of Europe. These responses act as a basis to this paper. Seven European countries have been chosen to be compared for this exercise- five of them are from the established democracies: Germany, the Netherlands, Austria, Sweden and France; and two are post-communist transitional justice/democratic countries: Poland and Moldova. These two sets of examples highlight the difference in focus between the two groups: in the first established democratic group: the focus is on the professional and qualitative standards, which include some focus on individual productivity; whereas the transitional democracies have a tendency to place more emphasis on productivity and quantitative criteria.

This is an appropriate place to highlight two issues further issues. Whilst it was requested that this speaker focus on bringing European experiences to the discussion taking place in Nigeria, two questions may arise: What is the relevance of performance evaluation systems for civil law countries to common law countries? And how would you approach the different professional attitudes of judges from Nigeria for evaluation differently than the European- especially the transitional democracies?

As to the first question: judicial performance evaluation has its roots in American practices (common law practices), and was designed, originally, to inform voters during retention elections of judges' performance in the last period. It is not within the scope of this presentation to discuss American experiences per se, but it can be said that the criteria used in America, expanded upon by the American Bar Association, and the National Centre for State Courts have provided a useful foundation for judicial performance evaluation in Europe.

As to the second question: given the versatile nature of the criteria and procedures used in performance evaluation of judges across the world, a solution can be found to suit the Nigerian context and professional approach: be it for self-improvement, improving overall productivity; improving public confidence; or discipline.

Comparative experiences²⁹

Rationale behind performance evaluation systems:

Judicial performance evaluation for the countries compared have usually focussed on evaluating aptitude, providing an objective system for promoting judges, and can be used in several of the countries as a system to dismiss judges who do not meet the standards expected of them (though this is still rare). Germany, the Netherlands, France, and Moldova for example explicitly state that the purpose of their evaluation systems are to assess performance and decide training needs of the judges, to foster 'quality of the judiciary'.

Sometimes promotion is of primary concern, sometimes of secondary concern: Germany, Poland, and Moldova expressly state that evaluations are to be used (though not exclusively) to provide an objective and merit basis for promotions; whilst evaluations can have consequences for promotions in Austria, France and the Netherlands. Poland is extremely ambitious, with a multi-layered evaluation programme, on the one hand for these two standard foci on aptitude and

²⁹ All comparative data for this part can be found at the following website: http://www.coe.int/t/dghl/cooperation/ccje/textes/Travaux17_en.asp Last accessed July 2 2014 Other data relating to Moldova has been taken from the OSCE report ""Assessment of the Performance Evaluation of Judges in Moldova", June 2014

improvement and promotions, and on the other hand it also has a separate evaluation programme for disciplinary proceedings; for part of appeals filed against entered judgments; and for violation of the right to be heard within a reasonable period of time (or not excessive length of time).

There are two anomalous countries: Austria and Sweden. Austria has no rationale behind its performance evaluation system in its law, but it can have consequences for promotions and may even lead to dismissal. This is anomalous as judicial independence is not adequately protected under a system that does not explicitly envisage having an impact on promotion or status of the judge. Sweden is anomalous because its aims are to provide a basis for an annual salary review, based on labour law rather than any separate legislative initiative for judicial reform. This also takes performance evaluation out of the context of judicial independence and into the practicalities of human resources.

The criteria

In seeking to achieve these various aims, these countries have a mixed approach of different methods, criteria, procedures and institutions. The questionnaire sent out by the Consultative Council of Judges in Europe gives lists of both quantitative³⁰ and qualitative³¹ criteria, and asks the respondents if they use them.

In the developed democratic countries, there are no set indicators for quantitative criteria as such, but statistics are used to gain an overall impression of judges' work. In Germany, it is not possible, because of judicial independence, to weigh judges' performance against any kind of absolute aim.³²

In the Netherlands, such data is used in the overall assessment of the judge. Where a judge fails 'to meet the average workload target, measures will be discussed to remedy the situation', including decisions of transfer within the rotation of the court, but in this situation, there is never a sanction or discipline- it is a matter of human resources- where best to place the judge that he/she can be most productive. Similarly in Sweden, they take account of quantitative data in so far as it assists the evaluator in understanding how the judge has met individually set goals in that year in order to assess the annual salary rise. In France and Austria, whilst they may take account of quantitative data, there no criteria for measurement as such. However, in the two transitioning democracies, Poland (member of EU), and Moldova (non-EU member) quantitative

³⁰ Under question 5: The number of cases in which a decision has been made by a judge; the average time spent on each of these cases; the average number of hearings per case; clearance rate (number of the cases, where a decision has been made, vis-à-vis the total of the cases forwarded to the judge); the average time to judgment (the time required to deliver a judgment by a judge after the completed hearing); any other quantitative indicators.

³¹ Under question 6: analysis of the type, subject and complexity of the cases dealt with by a judge and his/her decisions; the number of appeals vis-à-vis the number of the cases, where a decision has been made; the number of decisions reversed and/or cases remitted by the appellate court; the types of cases where decisions were reversed and/or cases remitted (criminal, civil, administrative or other); the grounds for reversal and/or remittal; any other qualitative indicators.

³² According to the German response to the CCJE questionnaire, an appeal is pending by a judge who's president reprimanded him for not meeting average performance- handling only 68% of the average number handled by other judges- but he claims he needs to take his time and apply other standards expected of him. This is likely to go to the Federal constitutional court.

indicators play an important role in the evaluation of judicial performance: Settlement rates and timeliness in light of the complexity of cases are observed in both countries.

The CCJE questionnaire also asks whether reversals are measured. Germany, Netherlands, Austria, Sweden and France do not measure reversals. As highlighted by the Kyiv recommendations, doing so creates tensions in the judiciary, leading first level courts to look over their shoulders, or worse, seek advice from courts of appeal on how to avoid reversals. This latter activity removes a level of jurisdiction and therefore a right to appellate justice. However, reversals play a role in evaluation in Poland and Moldova- which they say count towards the stability of case law (uniformity), but they have a different approach. Whilst the Polish do not look at grounds for reversal, Moldova examine reasons for reversal that are attributable to the judge (procedural and substantive errors). This also reflects the quality of decisions issued by individual judges.

Other qualitative indicators are important, as they measure the skills of the judge in social, organisational and professional contexts, an example of social skills which is an important aspect in the hearing of cases and treatment of litigants, lawyers and colleagues. Other indicators look at how trials are conducted, mediation capabilities, legal research, drafting skills, quality of motivation in decisions, and organisational skills (IT knowledge, organisation and professional activity). These indicators are measured to some extent in Germany, the Netherlands; Austria, France, Poland; and Moldova. What is important to emphasise is that several of these countries state explicitly that the content of a decision, especially in ongoing cases, can never be part of an evaluation. What is also significant is that, aside from Sweden, all of the other countries do measure ethical and professional issues (including the number of disciplinary offenses) and consider them to be an important aspect of performance evaluation of judges.

The Procedure

The body responsible for performance evaluation of judges varies between these countries: for some, the president of the court (court chair) is responsible, in other cases an independent body is responsible, and more rarely, the ministry of justice can also be responsible.

In Germany, the Netherlands, Sweden and France, it is a president of a court that is responsible. In Germany, the President of the Regional Court is responsible for all judges in his/her region, with a re-evaluation by the president of the court of appeal for balance and uniformity to ensure a lack of bias. Evaluation occurs every 4-5 years, with extraordinary evaluations possible where a judge seeks promotion. Here they have uniform evaluation forms to fill out with specific criteria to meet. In France, it is also the first president of the court of appeal for judges in their region

In the Netherlands the president of the court is responsible for evaluation of judges' performance. However, it is very informal and varies (in format) from one court to another. There is no single evaluation form (although guidelines are being currently being developed). Judges can participate during the interview as well as comment on the report. The informality is very important in a Dutch context and it has helped judges develop professionally under increasing spotlight on issues of productivity.

In Sweden, the head of the court of the first line manager (depending on size of the court) is responsible for performance evaluation for approving salary increases. This involves analysing activity of the previous year in light of the goals set out for the individual judge. Where the judge

has fulfilled the goals set out, the annual salary increase is going to be approved, with new goals set for the next year.

Therefore, court chairs in Germany, the Netherlands, France and Sweden hold a human resources role in developing and managing judicial capacity. Here there are fewer concerns about hierarchical control of judges whereby presidents of courts will dictate how judges should decide specific cases. Constitutional law and norms dictate that judges must be independent and impartial in delivering decisions in their cases, and this is respected by court chairs. There are other measures in place, such as appeals, media freedom, and complaints mechanisms to check these issues.

In Austria and Moldova however, it is not the court chair but an independent body that is responsible for evaluating judges. In Austria, there is the *personalsenat*, and the *rapporteur* is responsible for collecting reports and gathering data and the whole *personalsenat* is responsible for delivering a decision. In Moldova an independent evaluation board is responsible for deciding evaluation results. A *rapporteur* is assigned to each individual judge, for collecting reports and gathering data but the whole board is responsible for deciding a final grade together. This system shows an inherent lack of trust in court chairs to take responsibility for capacity building of their judges. It also shows more protection for judicial independence against hierarchical control within the courts.

Poland has a mixed system. The Poles use an inspector judge every four years, but at request of the president of the court. The board of the court analyses the conclusions, and may ask for further disciplinary proceedings against a judge. This creates external and independent accountability as well as internal checks and balances.

The type of data collected in these countries is quite comprehensive. As discussed earlier, in some countries, there is a focus on qualitative skills therefore observation of several different sources is important. In analysing quality of judgments, they need to examine cases decided by the judges (and in some cases, even at reversals). For comportment, observation of hearings is also important: the language used in communication with litigants, attitudes towards litigants, and even dress code in some places. Whilst emphasis may not always be placed on the quantitative aspects of a judges' performance, all countries do examine the statistical data on judges' performance, such as settlement rates and timeliness. Finally, the social skills of the judge are important too, and therefore comments from colleagues are also gathered (usually anonymously). Some courts and judges keep files about their own activities, such as taking part in training, or providing training, publishing, public functions and so forth.

Appeal

It is important in the performance evaluation of judges that an appeal is possible, given the consequences that results may have for the status of judges. This is to ensure that due process is respected, as well as judicial independence. In all countries it is possible to challenge or appeal a decision of an evaluation.

In Germany, an appeal is possible to two different courts: first to the administrative court and then to the judicial services court for the review of an evaluation decision. It may even be possible to appeal all the way to the federal constitutional court on a point of principle. In the Netherlands, an appeal is possible before the Central Court of Appeal. Decisions have been

challenged that resulted from evaluation, such as job rotation or promotions. In Austria, it is possible to appeal to the personalsenat of a court of the next higher level. In France, the decision can be contested before the Court of Cassation where the evaluation affects promotion; or to the Conseil d'Etat where the evaluator refuses to change the results of the evaluation. In Poland, there are also two different bodies to appeal to: firstly is the superior court that has responsibility for overseeing final results who can hear challenges during the evaluation itself; secondly judges may appeal to the Council of the Judiciary where a promotion opportunities are affected. In Moldova, results of evaluation may be appealed to the Superior Council of Magistracy, but only on issues of procedure and not substance.

The exceptional case is Sweden, where appeal is possible, but only to the Swedish Agency for Government Employers. This takes away the element of judicial independence, and focusses only on the professional aspect of judicial work and whether goals and objectives have been met. This is a very bureaucratic approach, and no report has been made as to its impact on judicial independence.

Consequences

Consequences vary between these countries for judges as a result of evaluation. The main consequence for judges is for promotion- this is the case for all countries discussed here except for Sweden, where the only consequence is for an annual salary rise. In the Netherlands, next to promotions, there can also be a consequence for job rotations within the court.³³ In Austria, Poland and Moldova it can also have consequences for discipline and dismissal (although this is rare and not recommended). In Moldova and Austria, judges need to receive insufficient grades in evaluations 2-3 times consecutively before dismissal proceedings may be started. In Moldova it is currently possible to dismiss a judge immediately for failing an evaluation (something that the OSCE have recommended that they change).³⁴

Sweden and Poland have an element of professional development as a consequence of its evaluation performance of judges. In Sweden, they set objective goals for a judge to meet (between the judge and the president) in order for a judge to receive an annual pay rise. In Poland they have individual career plans for self-improvement for each judge.

As stated in the theoretical framework, formal evaluation of performance for judges can provide an objective basis for promotion, either vertically (to a higher instance court) or horizontally (to a managerial position). The Venice Commission has stated that it is not ideal to base promotions purely on results of evaluations, and the process should also use interviews, motivation letters, recommendation letters and other factors. It is also recommended in the theoretical framework to keep discipline and dismissals as a separate procedure from evaluation in order to protect judicial independence.

³³ The Netherlands has a system whereby judges move between sectors of the court every 4-5 years (civil, criminal and administrative)

³⁴ "Assessment of the Performance Evaluation of Judges in Moldova", Organization for Security and Cooperation in Europe, Office for Democratic Institutions and Human Rights June 27 2014

Impact of Performance Evaluation of Judges

Comparative effectiveness

Comparative effectiveness can only be gaged depending on the rationale behind the performance evaluation system in place.

In Germany, an objective process for promotions has been established. According to the data provided to the CCJE, this is a better system for promotion than word of mouth. It is successful because the evaluator describes both positives and negatives, even though they know that their decisions may be appealed. It means that they deliver objective and independent evaluations. This is successful not only on the basis of its procedures, but also in appointing the right judges to the courts of appeal. In the Netherlands, their objective is different: theirs is about self-improvement and training. It is successful because it is achieved through informal means, whereby it is easier for judges to accept a need for change and improvement.

Whilst the Austrian's have not commented upon the success of their evaluation system for promotions, they have described another evaluation project where two judges are paired to observe the others' hearings in court. This has successfully led to better comportment towards litigants. The Swedish have also not commented upon the success of their own evaluation towards objective annual salary increases, it is therefore not possible to comment on the comparative success.

In France, performance evaluation has been criticised for not accurately reflecting the different circumstances and contexts of different courts and judges, and can therefore not be successful. Whilst they have been considering reform for a while, there is no consensus as to how to move forward.

In Poland and Moldova, both systems of performance evaluation of judges are in their early stages. However, in Poland, though its overall success is too early to assess, they have noted that the presidents of the court have not yet had time to develop individual career plans with the judges of their courts. In Moldova, in a report by the OSCE,³⁵ it was highlighted that they have failed to use it as an objective way to promote judges on the one hand, because the Superior Council of Magistracy may ignore the suggestions of the selection board. It has also failed in terms of skills and capacity building of judges, as the evaluations they have done are too superficial to be of any use to judges in self-improvement. There is also no connection to the training institute, therefore development of the expected skill set for judges must be done individually rather than institutionally. Furthermore, there has been a compromise to public confidence, as results of evaluations are published, leading to ranking of judges by grade, and litigants having less confidence in judges who have only received sufficient or below grades.

Comparatively therefore, the German and Dutch systems of performance evaluation of judges are the most successful in relation to their own goals and rationales.

³⁵ For which this speaker was an assessor and contributor

Impact on Productivity

As stated earlier, there is a mixture of approaches towards statistical data about productivity and settlement rates. However, having said that, one can see an impact on productivity in some of the countries.

In Germany for example, whilst they have not set a standard of productivity for judges to meet, the evaluator can highlight that a judge has not met a minimum standard of productivity, compared to the rest of his/her peers in the court.³⁶ The Netherlands, unlike Germany, has set a standard by which to measure judges' productivity, on the basis of their own averages (set out and reviewed on a regular basis). If standards are not met, organisational remedies are sought-but never sanctions, for example, it may be that a judge can be rotated to another section of the court where he/she may be more productive.

For the remaining countries, there can only be speculation as to whether there has been an impact on productivity. In Austria, for example, they do take account of quantitative criteria (though, like Germany, is not measured). The one complaint about the system is that too many judges are evaluated as 'excellent'. This may mean that performance regarding productivity, along with other criteria, is quite high. It could also mean that the standards have not been set high enough. Similarly in Sweden, one can assume that where judges receive their annual increase in salary, that they have met the objectives, both quantitative and qualitative.

In France, as they are in the middle of reform, one cannot make any conclusions about productivity. However, courts, like other public institutions, are funded under the organic law on finance (LOLF) which requires all institutions to meet certain objectives (management by objectives). For courts, this means "...delivering judgments of high quality within a reasonable amount of time in civil and criminal cases; increase and improve responses to criminal cases; manage the increasing costs of justice; and develop IT communications."³⁷ Of relevance to judicial performance on productivity would be the demand on timeliness. However, it is not possible, at this time, to say that judicial performance evaluation has had the desired impact of increasing productivity.

In Poland and Moldova it is too early to assess the impact of their relative systems on productivity of judges. Nevertheless, these are important criteria to both, and one can expect to see an improvement in productivity. Whether they can do so, keep the quality of judgments high, and in light of the requirements of judicial independence, is a question for later assessment.

Judicial Performance Evaluation in Nigeria

It is noted that, in several of the recommendations in the Annexure to the final report on reforms for the judiciary in Nigeria, there is a plan to implement a form of evaluation. It is not possible within this paper to go through all reforms that have been suggested by the sub-committees recommendations. However, it is possible to briefly highlight that performance evaluation systems can be a versatile tool to improve capacity in various areas.

Several issues of concern have been noted in the reforms that are appropriate for performance evaluation. The European systems above, for example, take into account issues of ethics and

³⁶ This is subject to an ongoing case in Germany at the moment

³⁷ Errera, Roger, *'Et ce sera justice... [And it will be justice...]*', Gallimard, 2013 p. 227 (own translation)

discipline in evaluations. Under paragraph 12 of the recommendations by the sub-committee on judicial integrity and dignity, it has been highlighted that “there is a need to appoint only persons of good character and integrity to the Bench.” This would be a standard to hold judges to during tenure, and can be re-evaluated.

In some states in America, during retention elections, courts publish results of evaluation performance of judges for the electorate to have a broader understanding of how well the judges have performed in the previous election period (this does not occur in Europe). In England, there is a system of performance evaluation of public defenders.³⁸ Such a system could be applied in Nigeria to ensure that judges being appointed to the Bench from the Bar meet certain professional, organisational, and ethical standards.³⁹

Another reform suggested by the Annexure is the ‘mandatory keeping of a court diary... by a judicial officer’ (paragraph 11 on the recommendations with regard to speedy dispensation of justice in practice and procedure in civil cases) as a move to increase transparency and increase public confidence. This would be measurable (where the technology is available) by an evaluation system, but also easy to keep track of by NGO’s such as Access to Justice that monitor court and judicial performance- and would go some way to enhancing public confidence.

Under observations on practice and procedure in criminal cases, evaluation of treatment of victims by judges can also be made to ensure that victims of crime are treated with respect and dignity throughout the process. This would be part of the qualitative criteria of professional skills and comportment (paragraph 5 improvement of criminal procedures).

Similarly, appropriate dress and comportment; perceptions of bias and political influence, hierarchical control within the judiciary can also be assessed and measured in evaluation systems, using qualitative criteria and indicators for appropriate comportment and compliance with the code of ethics (para 16-18 sub-committee recommendations independence).

In the recommendations found in the section by the sub-committee on constitutional amendments, independence of the judiciary and rebuilding public confidence, the sub-committee “expressed serious worry and concern over the drop in the standards of judgments, cases of conflicting judgments, and rising lack of respect for Legal Practitioners and colleagues by the Justices, all of which has led to a decline in confidence in the Court of Appeal” (paragraph 6). These are measurable indicators and are measured in the European countries compared in this paper (except Sweden). Similar standards can and should apply to judges of courts of appeal as to the lower court judges (with perhaps higher expectations on quality of judgments). Issues of uniformity are especially important for lower courts, and if courts of appeal have conflicting judgments, this can cause problems for lower courts’ uniform application of the law (paragraph 8). Professional behaviour of court of appeal judges can also be measured, and are measured in Europe (paragraph 11).

Several of the reforms recommended within the annexure to the final report, if implemented, can be easily measured for impact through an evaluation of both judges and courts (simultaneously).

³⁸ Bridges, Lee et al “Evaluation of the Public Defender Service in England and Wales”, 2007 pp.119-178 <http://www.law.cf.ac.uk/research/pubs/repository/1622.pdf> Last accessed 2 July 2014

³⁹ Paragraph 19 of the Annexure to the Final Report

Conclusions

To ensure a fairer and speedier delivery of justice is to ensure access to a fair trial that is operated by a competent and independent judicial body. Judicial performance evaluation is a limited tool of accountability- it is much more appropriate as a tool (in Europe) as a capacity builder, and can only focus on one aspect of a courts' work. The judge is not the only actor that can affect the outcome of cases. Judicial performance evaluation can help to improve the work of judges, and therefore of courts, but not independently of other reforms, which have been noted in the recommendations to the annexure of the final report on reforms for the judiciary in Nigeria.

The reform programme of Nigeria is comprehensive, complicated and will take many years of implementation before positive impact may be seen, especially for reforms that include the use of judicial performance evaluation. This is a useful reform tool as it can track, on a regular basis, the impact of any given reforms or recommendations. Examples from the USA and Europe tell us that performance evaluation systems are versatile and may be adapted to the contexts of a given country. With that in mind, and with the goals of reform in mind, Nigerian reformers would need to think carefully about the role that judicial performance evaluation should play in helping to ensure fairer and speedier delivery of justice.