
Syracuse Law and Civic Engagement Forum

Issue 4: On Judicial Independence

2016-2017

“PERSPECTIVES FROM THE BAR: Attorney Attitudes on the Implementation of a Federal Judicial Evaluation Program”

Peter Leasure ·
Lauren Bennett ·
Sarah McIntosh ·

The use of judicial performance evaluation is becoming more prevalent in the U.S. and in foreign jurisdictions. While such evaluations are indeed growing in prevalence, the federal judiciary has been largely resistant to such a program. The current study, through a survey of attorneys who have appeared before an Article III federal judge within the last three years, seeks to examine attitudes towards the implementation of a formal federal judicial performance evaluation program. Overall results indicate mixed attitudes towards implementing a federal performance evaluation program. Specific results for those for or against an evaluation program are discussed in detail.

· Peter Leasure, J.D. is a Ph.D. candidate in the Department of Criminology and Criminal Justice at the University of South Carolina. Direct all correspondence to Peter Leasure at pleasure@email.sc.edu.

· Lauren Bennett is a second year law student at Washington and Lee University.

· Sarah McIntosh is a first year law student at the University of South Carolina

**PERSPECTIVES FROM THE BAR: Attorney Attitudes on the Implementation of
a Federal Judicial Evaluation Program**

TABLE OF CONTENTS

I. INTRODUCTION

II. U.S. JPE PRACTICES

A. State

B. Local

C. Informal

D. Federal

III. FOREIGN JPE PRACTICES

IV. CURRENT STUDY and METHODS

V. RESULTS

VI. DISCUSSION

VII. CONCLUSION

I. INTRODUCTION

Judicial performance evaluation (JPE) programs are a relatively new phenomenon in which judges are evaluated on some set of criteria. Those in favor of such programs state the fact that judges, like every other profession, should be subjected to evaluation and given the opportunity to improve any shortcomings. It is stated that the benefits not only go to those in the court room including the judges themselves, but also to the public by increasing confidence in the judiciary. The use of JPE is not without opposition, though, as some state that such programs are problematic primarily because they constitute a threat to judicial independence, have too high a cost, and risk politicization. Despite this opposition, the use of JPE is becoming more prevalent as a large number of U.S. and foreign jurisdictions have successfully implemented various forms of JPE programs. However, the federal judiciary has remained largely resistant to such evaluations, which leads to several interesting questions. One such question asks whether attorneys who appear before federal judges support the implementation of a formal federal JPE program.

The current study seeks to answer this question by examining attitudes of attorneys towards the implementation of a formal federal JPE program. The current study will begin with a review of JPE programs at various jurisdictional levels within the U.S. and will also examine the use of JPE's in a few foreign jurisdictions. After this review, the current study will report the results of our survey. This paper will conclude by presenting any policy recommendations supported by the survey findings.

II. U.S. JPE PRACTICES

Currently, various forms of U.S. JPE programs are used in over 30 jurisdictions. The programs are to be discussed here in five broad categories: (1) State Programs; (2) Local Programs; (3) Informal Programs; and (4) Federal Programs. Most U.S. JPE programs are guided by five principles presented by the American Bar Association. The principles essentially provide benchmark measures for judicial evaluation. These five principles are legal knowledge, integrity and impartiality, communication skills, judicial temperament, and administrative skills.¹

A. State

There are essentially four types of JPE programs used at the state level.² The first type of state level JPE program evaluates judges for appointment or re-appointment and also self-improvement. For example, in Connecticut, the Judicial Performance Evaluation Advisory Panel, a permanent standing committee comprised of members of the bench, the bar, academia, and the Judicial Selection Commission, governs judicial evaluations.³ Attorneys can rate judges as excellent, good, fair, or poor for the categories of comportment, legal ability, and management skills.⁴ Jurors are given an opportunity to

¹ See American Bar Association Special Committee on Evaluation of Judicial Performance, Guidelines for the Evaluation of Judicial Performance (1985); American Bar Association, Black Letter Guidelines for the Evaluation of Judicial Performance (2005), available at, http://www.abanet.org/jd/lawyersconf/pdf/jpec_final.pdf.

² See Institute for the Advancement of the American Legal System (Judicial Performance Evaluation in the States), available at, <http://iaals.du.edu/quality-judges/judicial-performance-evaluation-states>, for a comprehensive list of states in each type of JPE program.

³ Connecticut General Statutes § 2-40a.

⁴ The District of Columbia evaluates judges on their work product, legal scholarship, dedication, efficiency, and demeanor. Retired judges seeking senior status are evaluated on their physical and

rate judges on issues of equality and fairness through a series of “attitude toward” questions and on other topics including dignity of proceedings, attentiveness, patience, courtesy, explanation of proceedings, efficiency, and clarity of charge.⁵ The evaluations are distributed semiannually and are analyzed once a reasonable sample size is reached.⁶ The evaluation results are distributed to the Judicial Selection commission and also to the Connecticut General Assembly’s Judiciary Committee and are not made public.

The second type of state level JPE program evaluates judges for election purposes. For example, Alaska was the first state to implement evaluations for the purpose of retention elections. In Alaska, the Alaska Judicial Council (AJC) governs the evaluations which are collected from all active and inactive members of the Alaska Bar Association, peace and probation officers in the state who handle state criminal cases, all court employees, jurors who have served with the judges, social workers, guardians ad litem, and child advocates. In addition, the AJC interviews the judges being evaluated, examines recusal and peremptory challenge records, and sends questionnaires to counsel who appeared before the judges under evaluation. The AJC also takes into account independent observations of judicial performance by the non-profit organization Alaska Judicial Observers, and holds public hearings on judges standing for retention to solicit citizen input. Judges are evaluated prior to election on legal ability, fairness, integrity,

mental fitness and ability to perform judicial duties. *See* Commission on Judicial Disabilities and Tenure, available at, <https://cjdt.dc.gov/page/evaluate-candidates>.

⁵ *Supra* note 3.

⁶ *See* New Jersey where mandatory evaluations are conducted twice during a judge’s initial term, with only the second evaluation going to the governor for reappointment. After reappointment, evaluation is voluntary. New Jersey Judicial Performance Program Rule 1.35A, available at, <http://www.njcourts.gov/attorneys/assets/rules/r1-35a.pdf>.

temperament, diligence and administrative skills. Results and a recommendation for each judge are included in the Lieutenant Governor's Official Election Pamphlet. Detailed evaluation results are posted on the AJC's website.⁷

The third type of state level JPE program evaluates judges simply for self-improvement. For example, in Massachusetts, judges must be evaluated at least once in a three year period on patience and attentiveness, preparedness, temperament, clarity of written and oral decisions, administrative capacity, control over the courtroom, legal knowledge, and fairness and impartiality.⁸ The evaluations are effectuated through surveys sent to attorneys, court employees, and jurors who appeared before the judge in the previous two years. The results of the survey are only for self-improvement and are therefore not released to the public.⁹

The fourth type of JPE program at the state level evaluates judges to improve public confidence in the judiciary. In these programs, results are compiled into aggregate numbers and individual judges are not identified. For example, in Hawaii, full-time judges are evaluated by attorneys and jurors on their legal ability, judicial management skills, comportment, settlement and/or plea bargain ability at the trial court level and on aspects of their fairness/impartiality, written opinions, and oral argument at the appellate level.¹⁰ Individual evaluation results are confidential and are provided only to the evaluated judge, the Chief Justice, and the Judicial Evaluation Review Panel. Individual

⁷ AS 22.15.195.

⁸ G. L. C. 211, § 26 and Rule 1:16 of the Rules of the Massachusetts Supreme Judicial Court.

⁹ *See also* Rhode Island who may release summary results to the public. Institute for the Advancement of the American Legal System, available at, http://iaals.du.edu/sites/default/files/documents/publications/rhode_islandjpe.pdf.

¹⁰ Hawaii Supreme Court Rule 19 (1993).

results are also available to the Judicial Selection Commission upon its request for appointment and retention purposes. A summary evaluation report for each court is made public.¹¹

B. Local

The local level programs generally follow procedures identified in the above state level programs. The only difference is that the evaluations are governed by a city level bar association. For example, in Ohio, the Columbus Bar Association annually evaluates on the basis of objectivity, judicial temperament, legal knowledge, quality of opinions, timeliness, and sentencing through surveys of attorneys who have appeared before a judge in the previous three years.¹² The results of the evaluations are made public for purposes of judicial elections.

C. Informal

The informal programs have the same goals as the above state and local level JPE programs (i.e. improving actual and perceived judicial performance); however, the informal programs are simply not supported by a formal organization or body. For example, the Washington State Joint Asian Judicial Evaluations Committee, including members from the Asian Bar Association of Washington, the Korean American Bar Association, the South Asian Bar Alliance of Washington, Filipino Lawyers of

¹¹ See also New Hampshire for a state using aggregate evaluations but with the option to release results to the public for judges who repeatedly scored poorly on evaluations. New Hampshire Supreme Court Rule 56.

¹² See Columbus Bar Association website, available at, http://www.cbalaw.org/CBA_PROD/Main/Resources/Resources_for_the_Public/Judicial_Election_Information/Main/Resources/Public/Judicial-Elections.aspx?hkey=6ea76aec-ed44-4539-b581-b24b77b8dbc0. For examples of other local programs, see the Houston Bar Association, available at, <http://www.hba.org/judicial-poll-results/>, the Austin Bar Association, available at, <http://www.austinbar.org/2015/02/austin-bar-releases-results-judicial-evaluation-poll/>, and the Dallas Bar Association, available at, <http://www.dallasbar.org/general-election-polls>.

Washington and the Vietnamese American Bar Association was created in part to ensure that judicial candidates are aware of and consider the issues important to the Asian Specialty Bars. Judges are evaluated on integrity, legal ability, fairness, and demonstrated commitment to equal justice.¹³ The evaluations are performed for election or appointment purposes where judges can be rated exceptionally well qualified, well qualified, qualified, or not qualified. This rating is valid for three years and is made public.

Another example of an informal program comes from Nevada where a local newspaper evaluates judges.¹⁴ Judges are evaluated on overall behavior and process oriented criteria by attorneys who have appeared before the respective judge. The evaluations are conducted every two years and are released by the newspaper.

D. Federal

The establishment of the Federal Judicial Center which was called “the research arm of the federal courts”¹⁵, the Civil Reform Act in 1990 which released semiannual reports to the public about federal case processing delays,¹⁶ and the Judicial Conduct and Disability Act¹⁷ which worked towards increasing judicial accountability through the creation of a formal complaint reviewing process all moved the federal judiciary in a

¹³ See Washington State Joint Asian Judicial Evaluation website, available at, <http://www.abaw.org/judicial-evaluations.html>.

¹⁴ See Las Vegas Review Journal website, available at, <http://www.reviewjournal.com/news/judicial-performance-evaluation>.

¹⁵ See Russell Wheeler, Empirical Research and the Politics of Judicial Administration: Creating the Federal Judicial Center, 51 LAW & CONTEMP. PROBS. 31, 38-39 (1988).

¹⁶ 28 U.S.C. § 476 (2008).

¹⁷ Judicial Conduct and Disability Act of 1980, 28 U.S.C. §§ 351-364.

positive direction in regards to improvement while on the bench. However, attempts to implement JPE programs at the federal level have been for the most part unsuccessful.¹⁸

In the 1970's, focusing on administrative law judges, several studies argued for the implementation of performance evaluations.¹⁹ For example, Administrative Conference of the United States Recommendation 78-2 stated as follows:

[m]aintaining the administrative law judges' decisional independence does not preclude the articulation of appropriate productivity norms or efforts to secure adherence to previously enunciated standards and policies underlying the [agency's] fulfillment of statutory duties.

Because of severe pushback (arguably from administrative law judges), implementation never occurred.²⁰ In 2006, a similar effort was made for evaluations of federal immigration judges; however, the implementation was once again unsuccessful.²¹

There have also been attempts to implement JPE programs for bankruptcy and magistrate judges at the federal level. The Seventh Circuit Judicial Council, for example,

¹⁸ See Rebecca L. Kourlis & Jordan Singer, *A Performance Evaluation Program for the Federal Judiciary*, 86 Denv. U.L.Rev. 7 (2008) for a review of federal judicial evaluation attempts.

¹⁹ General Accounting Office, *Administrative Law Process: Better Management is Needed* v-vi (1978); Administrative Conference of the United States Recommendation 78-2, *Procedures for Determining Social Security Disability Claims* 36 (1978); Administrative Conference of the United States Recommendation 86-7, *Case Management as a Tool for Agency Administration* 53 (1986).

²⁰ James P. Timony, *Performance Evaluation of Federal Administrative Law Judges*, 7 ADMIN. L.J. AM. U. 629 (1993); Jeffrey S. Lubbers, *The Federal Administrative Judiciary: Establishing an Appropriate System of Performance Evaluations for ALJs*, 7 ADMIN. L.J. AM. U. 589, 590 (1993); Toni M. Fine, *A Legislative Analysis of the Demise of the Administrative Conference of the United States*, 30 ARIZ. ST. L.J. 19, 59-61, 96-97 (1998).

²¹ See Press Release, U.S. Dep't of Justice, Attorney General Alberto R. Gonzales Outlines Reforms for Immigration Courts and Board of Immigration Appeals, Aug. 9, 2006; Nina Bernstein, *Immigration Judges Facing Performance Reviews*, N.Y. TIMES, Aug. 10, 2006.

has used evaluations to screen sitting bankruptcy judges who are applying for re-appointment.²² In that program, attorneys who appeared before a respective judge were randomly surveyed in an effort to make re-appointment decisions.²³

To the authors' knowledge, only two attempts have been made to implement JPE programs for federal district judges. The first was the Ninth Circuit in the 1980s. This program was modeled on two previous programs in California.²⁴ The Ninth Circuit Judicial Council ultimately adopted a voluntary, confidential self-evaluation program for district judges in 1981.²⁵ However, only nineteen of the two hundred thirty four judges eligible for the program—less than 8%—had actually undertaken self-evaluation.²⁶

The second effort to evaluate federal district judges came in the form of a pilot program completed in the Central District of Illinois in 1991.²⁷ In this program, the clerk of the court reviewed a pool of attorneys who had appeared in civil and criminal cases during the eighteen months prior to the study, and sent surveys to a sample of 150 selected attorneys who had appeared before each subject judge.²⁸ The results were only

²² Darlene R. Davis, *Judicial Evaluation Pilot Project of the Judicial Conference Committee and the Judicial Branch 2* (1991).

²³ *See also* the Eighth Circuit and Federal Judicial Center's programs for magistrate and bankruptcy judges. Federal Judicial Center, 2003 Annual Report 11; Davis, *supra* note 22.

²⁴ Kourlis & Singer, *supra* note 18.

²⁵ Davis, *supra* note 22.

²⁶ *Id.*

²⁷ *Id.*

²⁸ *Id.*

given to the judges under review and not made public.²⁹ Discussing the program, one judge stated that “The responses from the bar are an excellent barometer of how we are perceived to be performing our duties.”³⁰ Another judge stated that the results of the survey are “helpful because they are about as objective an evaluation as we can hope to get.”³¹ Despite the relative success of the pilot programs, a formal JPE program for federal article III judges has not been implemented.

There is also an informal federal judicial evaluation entity in the Almanac of the Federal judiciary.³² The Almanac provides judicial profiles of every federal and magistrate judges. These profiles are created from interviews with lawyers who have argued cases before the federal judiciary and various other demographic and case outcome information. Specifically, profiles include each judge's academic and professional background, experience on the bench, noteworthy rulings, notable media coverage, and commentary by lawyers about each judge's style, demeanor, knowledge, and management of courtroom proceedings. The Almanac is updated semi-annually.

III. FOREIGN JPE PRACTICES

Civil law countries regularly evaluate judges and have developed structured criteria for assessing judicial performance.³³ While civil law countries have somewhat different roles for their judiciary than common law systems, comparative analysis is still

²⁹ *Id.*

³⁰ Davis, *supra* note 22.

³¹ *Id.*

³² Almanac of the Federal Judiciary (Aspen Publishers 1995).

³³ Cheryl Thomas, *Report prepared for the UK Judicial Studies Board: Review of judicial training and education in other jurisdiction*, available at, http://www.ucl.ac.uk/laws/socio-legal/docs/Review_of_Judicial_Train.pdf (2006).

helpful to gauge the overall approval and approaches of JPE programs. France, Germany, and Austria are among the most highly established of these judicial evaluation programs. Austria requires judges to be evaluated by judicial boards every second year after the appointment to a new position.³⁴ The evaluations are done on the basis of several criteria such as legal knowledge, reliability, decision-making, work rate, conduct in high stress scenarios, communication, and social/personal behaviors.³⁵ The five possible grades for each criterion are excellent, very good, good, pass, fail.³⁶ If a judge receives a rating below “very good”, this warrants a follow up evaluation in a subsequent year. Any rating below “good” results in a pay cut for judges.³⁷

In France, evaluations are conducted every two years on an individual judge’s professional ability, legal and technical skills, organizational skills, and interpersonal skills.³⁸ The President of the Court of Appeal observes the judge, consults files, conducts an interview with the judge, and gains insight from other judges before publishing the final evaluation.³⁹ Evaluation results can lead to promotions, reassignments, or even transfers depending on the quality of skills observed. Judges can also challenge their evaluations.⁴⁰

³⁴ *Thomas, supra* note 33.

³⁵ *Id.* at 22.

³⁶ *Id.* at 125.

³⁷ *Id.*

³⁸ *Id.* at 116, 119.

³⁹ *Id.* at 33 at 119.

⁴⁰ *Id.* at 116.

In Germany, Section 26 of the Judiciary Act states that “judges are subject to service inspections only insofar as their independence remains unaffected.”⁴¹ Judicial evaluations are required every four to five years and also in cases of promotion.⁴² Evaluations are a requirement for all judges; however, judges over the age of 50 years are typically exempt from evaluation.⁴³ The President of Regional Court holds the responsibility to review and rank judges on professional, personal, and social competence and also competence to lead for more senior positions.⁴⁴ Specific factors include professional qualification, understanding of judicial office, ability to present persuasive arguments, ability to conduct hearings and interrogations, competence in teaching, personality, sense of duty and responsibility, workload cope-ability, management and organizational skills, decision-making, flexibility, and ability to work collaboratively and respectfully with others.⁴⁵ All of these elements are evaluated and ranked ranging from below average to excellent/very good and can lead to judicial promotions, transfers, or if not up to par, a mandatory leave of absence.⁴⁶

IV. CURRENT STUDY and METHODS

The current study seeks to examine attitudes of attorneys towards the implementation of a formal federal JPE program. This objective is achieved through a survey of attorneys who have been before an Article III judge at least once in the three

⁴¹ *Id.* at 125.

⁴² *Thomas, supra* note 33 at 121.

⁴³ *Id.*

⁴⁴ *Id.*

⁴⁵ *Id.* at 123.

⁴⁶ *Id.* at 121.

years before receiving the survey.⁴⁷ A sampling frame of participants was created by first randomly selecting a federal circuit. The federal circuit randomly selected was the Sixth Circuit which includes Ohio, Michigan, Kentucky, and Tennessee. Member lists for each of those states bar associations⁴⁸ plus the federal bar association member list were then compiled wherefrom a random sample was then drawn.⁴⁹ The 6th Circuit is an excellent jurisdiction for study as those states possess varying methods of judicial selection and various stances on JPE programs.

Tennessee used to utilize a JPE program; however, it expired in 2014.⁵⁰ The program was used to recommend the retention of appellate level judges.⁵¹ Tennessee selects trial court judges via partisan elections and appellate court judges via merit selection with retention elections.⁵²

Michigan has a formal JPE program⁵³; however, the program seems to be rather limited compared to other JPE programs and focuses on issues such as case processing times. Judges are selected in non-partisan elections in Michigan.⁵⁴

⁴⁷ A first question not included in the analysis below asked “Have you been before an Article III federal judge within the last three years?”

⁴⁸ See OHIO BAR, available at, <https://www.ohio-bar.org/Pages/Find-A-Lawyer.aspx>.

⁴⁹ For Tennessee and Michigan, the largest aggregate search term was a city. Samples were drawn from large cities such as Nashville and Memphis for Tennessee and Detroit for Michigan.

⁵⁰ See TENNESSEE STATE COURTS, available at, <http://www.tncourts.gov/boards-commissions/boards-commissions/judicial-performance-evaluation-commission>.

⁵¹ *Id.*

⁵² See JUDICIAL SELECTION, available at, http://www.judicialselection.us/judicial_selection/index.cfm?state=TN.

⁵³ Mich. Comp. Laws § 600.238 (1996).

⁵⁴ See JUDICIAL SELECTION, available at, http://www.judicialselection.us/judicial_selection/index.cfm?state=MI.

Ohio utilizes a JPE program at the local level via a local bar association.⁵⁵ As to the method of judicial selection, Ohio uses a non-partisan election system.⁵⁶

Kentucky, like Ohio, has a local JPE program also via a local bar association.⁵⁷ The program considers a number of factors making it similar to the more robust programs detailed above. Kentucky selects judges via non-partisan elections.⁵⁸

As survey response rates have been low for courtroom actors in the past, the survey instrument was kept as short as possible in hopes to increase the response rate. Initially, 500 surveys were sent out to the 6th Circuit states and the federal bar association members. Unfortunately, response rates were very low. In order to increase sample size, responses collected in the pilot phase (sent out randomly and containing the final version of the survey instrument) were also included in the analysis.⁵⁹ In the end, 600 surveys were sent out and 186 were returned,⁶⁰ resulting in a 31% response rate.⁶¹

⁵⁵ See Columbus Bar Association, available at, http://www.cbalaw.org/CBA_PROD/Main/Resources/Resources_for_the_Public/Judicial_Election_Information/Main/Resources/Public/Judicial-Elections.aspx?hkey=6ea76aec-ed44-4539-b581-b24b77b8dbc0.

⁵⁶ See JUDICIAL SELECTION, *supra* note 54.

⁵⁷ See Louisville Bar Association, available at, [https://www.loubar.org/UserFiles/files/Judicial%20Evaluation%20Archives/2015/2015%20Judicial%20Evaluation%20Comparative%20\(d3%2011-09-15\)%20FINAL.pdf](https://www.loubar.org/UserFiles/files/Judicial%20Evaluation%20Archives/2015/2015%20Judicial%20Evaluation%20Comparative%20(d3%2011-09-15)%20FINAL.pdf)

⁵⁸ See JUDICIAL SELECTION, available at, http://www.judicialselection.us/judicial_selection/index.cfm?state=KY.

⁵⁹ Pilot surveys were sent to various states located outside the 6th Circuit. However, we first analyzed the results for the pilot surveys and the results for the 6th Circuit/federal bar separately to see if there were any important differences. There were no differences detected, therefore results are presented as overall findings.

⁶⁰ Surveys returned by jurisdiction are as follows: 31 from Ohio, 27 from Kentucky, 22 from Tennessee, 31 from Michigan, 26 from the Federal Bar, and 49 from the multiple pilot states.

To analyze data, pattern matching was the guiding technique.⁶² During initial coding, memoing and theoretical sampling, where data is brought to codes and codes to categories, was utilized.⁶³ In summary, two levels of analysis took place: (1) open coding and (2) organizing the codes into categories or axial coding. To ensure reliability and validity, the study utilized peer examination of the data, where multiple researchers independently come to consensus on the emerging patterns, which helps to report consistent findings.⁶⁴

V. RESULTS

QUESTION 1. Do you think we should have formal federal judicial performance evaluations in your circuit? Why or why not?

As to the initial question, 109 (59%) answered yes and 77 (41%) answered no. For those who answered yes, the most common pattern (83) detected stated that judges should be accountable because most other professions are not immune from evaluation. Several respondents provided excellent examples of this pattern:

R22: *Everyone needs to be held accountable.*

R13: *I think this provides a measure of accountability, even for life-tenured judges.*

R134: *In every job you are evaluated, judges should not be an exception. I especially support evaluations of anybody in a position of responsibility.*

⁶¹ Such a response rate has been deemed acceptable in several other studies within the legal field. See Alec Ewald & Marnie Smith, *Collateral Consequences of Criminal Convictions in American Courts: The View from the State Bench*, JUSTICE SYSTEM JOURNAL, 29(2), 145-165 (2008).

⁶² Robert K. Yin, *Qualitative Research From Start to Finish*, Guilford Publications, (2015).

⁶³ Kathy Charmaz, *Constructing Grounded Theory: A practical guide through qualitative research*. Sage Publications Ltd, London (2006).

⁶⁴ Laura Krefting, *Rigor in qualitative research: The assessment of trustworthiness*, AMERICAN JOURNAL OF OCCUPATIONAL THERAPY, 45, no. 3, 214-222, (1991).

The second most common pattern (64) stated that evaluations help secure high quality judges. Examples of individual responses were as follows:

R24: Having periodic evaluations would show the judge in which area(s) he/she needed to improve.

R67: Judges need feedback to see how they are being perceived by those who practice and appear before them.

R17: We need evaluations, but we have to be careful that any evaluation remains objective, and that the bases for the evaluation does not inject any criteria that could be used politically. The reasons for evaluation are (1) to have some objective criteria on performance, (2) to provide an incentive for federal judges to continue to provide thoughtful, quality, timely and considerate rulings, and (3) to provide a basis to determine whether a judge remains competent to continue to preside. Currently, federal district court judges have very little oversight with respect to quality.

R177: Most judges who are appointed to the federal bench have never played that role before getting their commissions. Although they attend judges school at the Federal Judicial Center, they are hardly masters of their domains. Many have never been trial lawyers and don't have a mastery of the rules of evidence and procedure. Evaluations of the judges, albeit anonymous, could be valuable teaching tools that would enable them to improve their craft as they gain more experience.

R111: As a practical and procedural matter, federal magistrates handle much of the preliminary matters before the case reaches the Judge that will try the case. [This means that the article III judges] do not get "feedback" from witnesses, lawyers and jurors as do state court judges.

The third most common pattern (39) detected stated that evaluations would enhance public confidence in the judiciary. Examples of this pattern as are follows:

R22: Evaluations add legitimacy to the courts

R54: Judges should be held to very high standards to ensure there is no perception of corruption or ineptitude.

R78: Courts are the vessel by which we solve disputes between people. Judges are the public face of that process and evaluations enhance the public's view that the judiciary is competent

The final pattern detected (28) stated that the current mechanisms in place simply do not do enough to ensure the best possible judiciary. Respondents stated as follows:

R15: Evaluations could help us detect small problems that may be occurring now such as a personal problem or a decrease in accuracy of records or timeliness before they become a larger problem that can affect other areas

R88: A judge appointed for life may be a terrible judge who never commits a personal transgression such that he or she loses their bench. So this type of evaluation may be one of the only ways to correct smaller issues

For those who answered no, the most common pattern detected stated that evaluations would be too costly (55). One respondent stated as follows:

R55: The extra resources required to do the special evaluation is unwarranted. Our judges know above all to uphold the integrity and independence of the judiciary, and to avoid impropriety and appearance of impropriety in all of their activities

The second most common pattern stated that evaluations would not be effective in changing judicial practices (49). One respondent noting the criminal aspect of judicial proceedings stated as follows:

R91: I think it is really the prosecutors should be evaluated. Prosecutors have the true discretionary power in dismissing a case, plea bargaining, and recommending sentencing to judges. Why should there be even further scrutiny of judges when judges feel incapacitated to be able to dismiss cases or veer away from recommended sentencing?

Also within this pattern, another respondent noted life-time tenure as a reason not to do evaluations:

R66: These judges have lifetime tenure and evaluations would be a waste of time. The work of a judge cannot be judged by metrics, as much as that is in vogue right now.

Another pattern (45) stated that evaluations would harm judicial independence.

Respondents stated as follows:

R179: I think evaluations of judges would be a dumb idea, one that would inevitably --- no matter who designs the criteria, no matter who conducts the evaluations ---lead to political interference by various interests groups that would result in influencing judges' decisions and the outcomes of cases. Popularity of rulings/outcomes would end up trumping legal soundness as the template for judicial quality. A mischievous idea that would diminish, not enhance, the quality of judging, an idea that ought to be buried without delay. Better idea: make sure the initial selection of judges is done with thorough screening and vetting for integrity, intelligence, writing ability, temperament, courage and experience.

R2: I believe judicial evaluations compromise the independence of the judiciary. Any results would be politicized and subject to lobbying efforts, whether liberal or conservative. It's best to stay out of the evaluation game.

R69: Evaluations would have a chilling effect on judicial discretion, and could/would be subject to manipulation by the evaluators

R7: Our Constitution allows judges to have life-terms for a reason – so that the federal judges are not pressured by politics, lobbying, and other non-legal pressures. If we introduce external, mandatory evaluations to the work of the judiciary, then we will be introducing an external influence into the judicial process.

Another pattern (29) stated that evaluations would be subject to bias. For this pattern, respondents raised questions about who would do the evaluating and on what items judges would be evaluated.

R7: It seems to me like evaluations could be biased depending on who fills out the paperwork. I'm sorry I am just not a fan of these.

A final pattern (28) stated that evaluations were unnecessary because current processes ensured high quality judges.

R180: I think we already have a sufficient system in place for evaluating judges because of appellate review. A judge that acts beyond the scope of authority or in an arbitrary manner can be reversed by a higher court reviewing the decision.

R34: I would object to any proposal to evaluate judges, state or federal, during their tenure. Why? The legal and professorial community evaluate judges daily and spread the word as to the good, the bad, the ugly.

R182: I think a strong selection process of judges would eliminate the need for judicial evaluations

QUESTION 2. If you think there should be evaluations, on what should judges be evaluated?

There were 103 responses to this question and each respondent noted at least one suggestion and most noted several. Responses did not raise any criteria that was not currently in use in one of the jurisdictions currently using JPE's. The most common response was professionalism/courtroom demeanor (88). Another common pattern (72) was knowledge of the law/competency. Sixty six respondents suggested that judges should be evaluated based on impartiality. The ability to efficiently manage a docket/issue decisions had 65 responses. Forty seven respondents noted that reversals on appeal should be a criterion for evaluation. Availability to parties/communication had 46 responses and 33 respondents noted that number of conduct complaints should be included as well.

QUESTION 3. If you think there should be evaluations, who should do the evaluations?

There were 104 replies to this question. The most common response (71) stated that attorneys who appeared before the respective judge should do the evaluation. One respondent stated as follows:

R56: Evaluations from peers such as attorneys may often have greater impact on reforming a bad judge's behavior and encouraging good judges to remain good.

An appointed neutral and detached committee was the next most common response (69). Example responses are as follows:

R76: Independent, multi-partisan committees of both lawyers and non-lawyers (which should include ordinary citizens - not solely business and corporate leaders), including retired judges. And small enough body to function successfully but large enough to be representative.

R99: An appointed body that has lawyers & non-lawyers, gender, ethnically & racially diverse, diversity in terms of income and employment, etc.

R13: It would need to be an independent arm of the federal judiciary, like an IG or GAO

The final pattern (47) stated that there should be a mix of evaluators. Examples of this pattern are as follows:

R186: 360 degree evaluations where everyone who has official contact with judges except for plaintiffs and defendants (they're going to be biased, e.g., you don't want criminals sentenced to prison to rate their judges). Instead, court reporters, counsel who have appeared before judges, security, and other judges.

R14: Anyone that works with them regularly, similarly to how they do peer evaluations in other work environments.

QUESTION 4. If you think there should be evaluations, should the evaluation results be made public?

Of the overall 103 responses for this question, 63 said yes and 40 said no. For those in favor of making the evaluations public, the most common response (47) stated that judges may be encouraged to improve upon evaluations if they know the public will see results.

R55: It can be really helpful to improve the justice system if the judges know the public will see their evaluations.

Another pattern (46) for those in favor of making the evaluations public stated that the public simply has a right to know of the results. For example, respondents stated as follows:

R3: Ultimately judges are there for the community. The community should be able to know the performance of those justices that ultimately preside over them.

For those in favor of making the results of the evaluations private, the most common pattern (33) stated concerns over the public not possessing enough knowledge of the judicial process to accurately interpret evaluation results.

R25: I do not think the evaluations should be made public because it may affect the independence of the judiciary by compromising the judge's ability to make unpopular but legally correct decisions. If the goal is to improve a judge's performance, then there is little value in exposing weaknesses to the public as a means of achieving that goal. Improvement would be achieved through a carefully devised plan that focused on areas where a judge did not display sufficient competency.

Another common response (27) stated that making evaluations public is unfair to judges because releasing findings of an evaluation could lead to a negative view of a particular judge without their ever having the opportunity to fix the issue. Example responses are as follows:

R166: Judges, particularly those just beginning their careers on the bench, should have the opportunity to improve and therefore more effectively serve the public interest. The purpose of the evaluations is to ensure accountability, not to terminate judges based on a single evaluation done by a finite number of people.

R5: If the evaluation is used to catch small problems before they evolve, letting this information into the public sets those problems in stone when they could still be corrected.

Another small pattern (9) emerged that stated results could be posted in the aggregate without names. This is consistent with some state JPE practices.⁶⁵

QUESTION 5. If you think there should be evaluations, how often should they be done?

There were 97 responses to this question. The responses to question five did not produce any results that were dissimilar to common state practices. Therefore, overall patterns will only be discussed. However, a few patterns are worthy of note. One pattern detected (13) stated that judges who received poor evaluations in the past should be subject to more frequent evaluations. Another pattern (7) stated that every case should have an evaluation procedure.

R181: Every case--judges at the district court and the Sixth Circuit Court of Appeals have the ability to evaluate the performance of the appointed counsel who appear before them and there is a pre-printed evaluation that can be used. Nothing wrong with a similar vehicle here--it can be deposited in the clerk's office and anonymity can be maintained.

QUESTION 6. If you think there should be evaluations, should a judge who does poorly face some sort of sanction?

There were 106 responses to this question. Seventy six respondents believed there should be some type of sanction for judges who do poorly on evaluations. Various types of sanctions were recommended. The most common pattern (64) stated that sanctions should be geared more toward self-improvement and education or counseling.

Respondents stated as follows:

R56: I think they should concentrate more on CLE or CJE in the judge's case. Continuing Judicial Education.

R19: I really think this should be a thing for self-improvement.

⁶⁵ See New Hampshire, *supra* note 11.

Forty nine respondents recommended a graduated sanctioning system.

Respondents stated as follows:

R11: *The judge should first be confronted about the issue and given options on how it could be resolved. If the action were to continue, some sort of reprimand would be necessary.*

R9: *There should be a graduated system: first a warning, then reprimand, etc.*

Finally, some respondents (17) noted that judges who score poorly in certain areas could be removed from certain case types and thirty respondents did not believe that sanction would be necessary after from a poor evaluation.⁶⁶ For example, one respondent stated as follows:

R98: *It may keep good candidates from accepting the post.*

VI. DISCUSSION

The above results present several interesting findings. A majority of respondents noted that they would favor a formal federal JPE program. The most prominent reasons for the implementation of a formal federal JPE program were accountability⁶⁷ and improved quality of judges, while some also noted that evaluations would increase the public's positive view of the federal judiciary.

For those who were not in favor of implementing a federal JPE program, the most prominent responses were that evaluations would harm judicial independence, were too

⁶⁶ Nine respondents were unsure about sanctions.

⁶⁷ See Kourlis & Singer, *supra* note 18 and David C. Brody, *Use of Judicial Performance Evaluation to Enhance Judicial Accountability, Judicial Independence, and Public Trust*, DENV. U.L. REV., 86, 115 (2008) for a further discussion of the benefits of evaluations on accountability.

costly, and would not likely be effective in changing judicial practices. On the judicial independence finding, it is important to note that there have been empirical studies that sought to gauge the effect of JPE programs on judicial independence.⁶⁸ A 2008 survey of Colorado judges found that 28% of trial judges felt that the state's JPE program decreased their judicial independence. Forty four stated that their JPE program had no effect and 29% stated that the program in fact increased their independence.⁶⁹ In another study, only 14.5% of judges in Colorado, 22% of judges in Alaska, and 33% of judges in Arizona indicated their belief that independence was undermined by their state's JPE program.⁷⁰

Another response from those not in favor of a federal JPE program stated that such evaluations have too high a likelihood of bias to justify their implementation. On this point, there is recent research that indicates that bias is indeed a problem with JPE programs.⁷¹ In that study, it was found that implicit bias (a theory of social cognition) likely resulted in minority and women judges consistently scoring lower on state JPE programs.⁷² In response to these findings, one author stated that JPE programs "are the best tool we currently have to capture public perceptions of judicial competence,

⁶⁸ See also Penny J. White, *Judging Judges: Securing Judicial Independence by Use of Judicial Performance Evaluations*, FORDHAM URB. LJ, 29, 1053 (2001).

⁶⁹ Institute for the Advancement of the American Legal System, *The Bench Speaks on Judicial Performance Evaluation: A Survey of Colorado Judges* (March 19, 2009), available at, <http://ssrn.com/abstract=1365256> or <http://dx.doi.org/10.2139/ssrn.1365256>.

⁷⁰ Kevin M. Esterling & Kathleen Sampson, *Judicial Retention Evaluation Programs in Four States: A Report with Recommendations*, American Judicature Society (1998).

⁷¹ Rebecca Gill, *Implicit Bias In Judicial Performance Evaluations: We Must Do Better Than This*, JUSTICE SYSTEM JOURNAL, 35(3), 301-324 (2014).

⁷² *Id.*

professionalism, and commitment to procedural fairness—the same perceptions that form the foundation of the courts’ institutional legitimacy. Surveys also give judges special insight into how they are viewed by the lawyers who appear before them—views that judges are otherwise unlikely to receive.”⁷³ The author also proposed several measures which could reduce implicit bias in evaluations results.⁷⁴

As to what measures to use in evaluations, respondent answers were fairly well aligned with much of the current state and local measures of judicial performance. Patience, attentiveness, preparedness, temperament, clarity of written and oral decisions, administrative capacity, control over the courtroom, legal knowledge, and fairness and impartiality were all common measures that were noted by respondents. Responses on how often to perform the evaluations were also consistent with state and local JPE practices. As to who should perform the evaluations, the most prominent answers were attorneys who appeared before a particular judge, independent detached committees, and anyone who comes into regular contact with a particular judge (i.e. lawyers, jurors, court personnel, parties, other judges).

As to whether the evaluations should be made public, prominent responses in favor of making evaluations public stated that releasing results to the public would motivate judges to improve upon shortcomings. Similarly, respondents also noted that judges are ultimately public servants and that the public has a right to be made aware of any performance results. For those not in favor of making evaluations public, prominent responses argued that releasing the evaluations would harm judicial independence.

⁷³ Jordan M. Singer, *Attorney Surveys of Judicial Performance*, JUDICATURE, 98, 20 (2014).

⁷⁴ *Id.*

Another interesting response from those against making results public was that it was unfair to release negative results of a particular judge before they had a chance to rectify any shortcomings.

As to implementing some sort of sanction for a poor evaluation, those in favor of doing so stated that such sanctions would give a JPE program teeth. Sanction recommendations mostly involved internal or public reprimands. Those against some sort of sanction for poor evaluation results stated that sanctions would be unnecessary given life-tenure.

VII. CONCLUSION

While judicial performance evaluation systems are indeed growing in prevalence, the federal judiciary has been largely resistant to such a program. Our study sought to examine attorney attitudes towards the implementation of a formal federal judicial performance evaluation program. Overall, our results indicate mixed attitudes towards implementing a federal performance evaluation program. If the federal system or any other jurisdiction decides to explore the implementation of a judicial evaluation program, our review of JPE programs and current findings provide a plethora of viewpoints on what such a system should entail. Weighing all the above viewpoints presented in this paper, we believe the benefits of a federal evaluation system outweigh any negatives. Consistent with this position, we think it is best to quote the Honorable Richard A. Posner. Judge Posner stated as follows in regards to judicial evaluation and development:

Judges would benefit from praise that indicated where they were doing a good job as well as from criticism, and judges who were not praised would learn from the praise of others where they were falling short.⁷⁵

⁷⁵ Richard A. Posner, *How Judges Think*, Harvard University Press, 204, (2010).