



NATIONAL PERSPECTIVES ON JUDICIAL PERFORMANCE EVALUATION

CONSISTENCY • TRANSPARENCY • HONESTY • IMPARTIALITY • INTEGRITY • ACCOUNTABILITY • EFFICIENCY



INSTITUTE *for the* ADVANCEMENT
of the AMERICAN LEGAL SYSTEM



Most Americans undergo job evaluations, and there is no reason why judges should not do the same. Judicial performance evaluation (JPE) processes were first developed in the 1980s and provide a foundation for states to assess the job performance of judges. Today, JPE programs continue to focus on the right goal of holding judges accountable for job performance based on politically neutral qualities like impartiality, transparency, and consistency rather than their rulings in cases. However, updating evaluation methods is essential to keeping JPE relevant and useful. Current JPE processes no longer fully capture the experience of modern court users, the needs of modern judges, or the expectations of modern voters. IAALS' JPE 2.0 project will help JPE programs update their approaches to reflect modern realities, while remaining accurate, trusted, and relevant.

Learn more at iaals.du.edu/jpe.

JPE 2.0

NATIONAL PERSPECTIVES

ON JUDICIAL PERFORMANCE EVALUATION

Danielle Kalil

Director of Civil Justice and the Judiciary

Logan Cornett

Director of Research

James Swearingen

Research Analyst

July 2024

For reprint permission, please contact IAALS.

**Copyright © 2024 IAALS, the Institute for the
Advancement of the American Legal System.**

All rights reserved.



INSTITUTE *for the* ADVANCEMENT
of the AMERICAN LEGAL SYSTEM



IAALS—Institute for the Advancement of the American Legal System

John Moye Hall, 2060 South Gaylord Way, Denver, CO 80208

Phone: 303-871-6600

iaals.du.edu

IAALS, the Institute for the Advancement of the American Legal System, is a national, independent research organization that innovates and advances solutions that make our civil justice system more just.

Founded in 2006 at the University of Denver, IAALS believes that justice for all must be a reality for everyone. When innovation is rooted in finding common ground, questioning the status quo, and centering the people, we begin to craft solutions that transform our civil justice system. IAALS' unique approach depends on purposeful research, deep collaboration, and diversity of perspective, followed by evidence-based recommendations that take hold in courts and legal institutions across the country—jumpstarting the groundbreaking and achievable solutions that will clear a path to justice for everyone. Because justice for all will never be a reality if those seeking justice cannot access the system designed to deliver it.

Brittany K.T. Kauffman	CEO
Danielle Kalil	Director of Civil Justice and the Judiciary
Janet Drobinske	Senior Executive Legal Assistant
Jordan M. Singer	Consultant on the Judiciary
Logan Cornett	Director of Research
James Swearingen	Research Analyst

TABLE OF CONTENTS

Introduction.....	1
JPE Programs in Study States.....	3
Methodology.....	4
Results.....	5
Perspectives on the JPE Process.....	6
Perspectives on JPE Evaluation Tools.....	9
Perspectives on JPE Evaluation Results.....	15
Perspectives on Commissions.....	19
Most Important Qualities in a Judge.....	20
Discussion and Conclusion.....	21

ACKNOWLEDGEMENTS

We would like to thank the judges who took the time to complete this survey and provide their perspectives on JPE, without whom we would not have been able to gather the data presented in this report. We are also grateful for the expertise our partners provided as we developed, distributed, and reviewed the survey. This report would not have been possible without the assistance of those in Alaska, Colorado, Hawaii, Idaho, Massachusetts, New Mexico, Utah, and Virginia who helped us distribute this report to judges. Finally, we are very thankful for the contributions of current and former members of the IAALS JPE 2.0 Task Force and IAALS JPE Working Group who facilitated distribution of the survey in their states and provided feedback on the report. It would not have been possible to provide the data contained here without the invaluable contributions of the many individuals who participated in this effort:

Barbara Arnold, Former Program Manager, New Mexico Judicial Performance Evaluation Commission

Adam Cohen, Planning and Program Evaluation Division, Hawaii State Judiciary

Susanne DiPietro, Executive Director, Alaska Judicial Council

Anna Rachel Dray-Siegel, Assistant Legal Counsel, Massachusetts Supreme Judicial Court

Farrah Fite, Former Communications Director, The Missouri Bar and Staff Liaison,
Missouri Judicial Performance Review Committee

Michael Oki, Research Statistician, Hawaii State Judiciary

Mateo Page, Former Director, Court Services Division, New Mexico Administrative Office of the Courts

Mary-Margaret Pingree, Executive Director, Utah Judicial Performance Evaluation Commission

Andrea Powers, Human Resources Director, Idaho Judicial Branch

Jordan Singer, Professor of Law, New England Law | Boston

Kent Wagner, Executive Director, Colorado Office of Judicial Performance Evaluation

Kristi Wright, Program Director, Virginia Judicial Performance Evaluation Program

Jennifer Yim, Former Executive Director, Utah Judicial Performance Evaluation Commission

INTRODUCTION

Judicial performance evaluation (JPE) processes—referred to as judicial performance review (JPR) in some jurisdictions¹—provide a foundation for states to assess the job performance of judges. JPE programs were first developed in the 1970s and, currently, 16 states², the District of Columbia, and Puerto Rico have implemented official JPE programs. These programs continue to focus on the goals of evaluating judges for politically neutral qualities like impartiality, clarity of communication, and legal knowledge, along with accountability for job performance rather than the specific decisions the judge has made.³

JPE program objectives may differ depending on the state. All programs share the internal goal of fostering judges' professional development and self-improvement. In some states, these programs are also designed for the external goal of informing elected officials or the public about the judiciary's performance. Finally, in other states, JPE is used to inform those making decisions about whether to retain judges, whether that be voters, the state legislature, or another government body.⁴

All of these goals are encompassed by the broader objective of promoting judicial accountability and transparency while maintaining judicial independence.⁵ When they are working well, JPE programs are meant to promote public trust and confidence in the judiciary and our court system.⁶ They offer a mechanism for holding judges to high standards of performance and provide an opportunity for judges to improve how they are serving the public. When the public trusts that judges are doing their jobs fairly and efficiently, they are more likely to seek help from the courts and comply with their rulings.⁷ However, updating the evaluation approaches used in JPE programs is essential to ensuring that they reflect the experiences of modern court users, the needs of modern judges, and the expectations of modern voters.

Despite their critical importance, JPE programs have suffered from waning enthusiasm in the past decade, with some critics expressing concern about the accuracy, objectivity, and validity of evaluations and the lack of overall transparency in the process.⁸ Additionally, changes in society and our courts—for example, the use of technology and the changing representation of litigants⁹—have altered the public's perception of and relationship with today's courts. JPE programs must adapt. Even longstanding and well-regarded JPE programs need to update their approaches to remain accurate, trusted, and relevant. As is demonstrated in the results of this study, while judges generally believe JPE programs are important, they identify many opportunities for improvement.

IAALS, the Institute for the Advancement of the American Legal System at the University of Denver, has long been at the forefront of policy innovation related to JPE, convening stakeholders from across the country and working directly with states to improve and advance their programs. IAALS has hosted the National JPE Working Group since 2007, which brings together state JPE coordinators quarterly to exchange information and ideas. Judicial performance evaluation remains a critical component of IAALS' recommendations for promoting effective courts that merit public trust. As part of our *O'Connor Judicial Selection Plan*, we recommend that judges be evaluated near the end of their term by diverse groups of people who have knowledge of their performance, including attorneys, litigants, jurors, witnesses, court employees, law enforcement officials, and victims.¹⁰ We then recommend that a JPE commission review survey results, rate a judge's performance, determine whether or not they should be recommended for retention, and make the summary of their findings available to the public before a retention election. IAALS has published a number of recommendations for JPE programs over the years, using what we have learned from our research and our convening of stakeholders to issue the most up to date best practices for evaluating judges.¹¹

The most recent iteration of our work in this area is our JPE 2.0 project, which was launched to develop new best practices for modernizing JPE, thinking creatively about how to maintain its core goals while also being responsive to emerging techniques and legitimate concerns about historic methods. JPE 2.0 is a multi-stage effort, including:

- **Comprehensive background research** on the history of JPE programs, current perspectives, and issues warranting focused consideration
- **Collaboration** with IAALS' JPE 2.0 Task Force to get input and expertise from administrators and experts around the country
- **A national in-person convening**, JPE 2.0: Modernizing Judicial Performance Evaluation, to identify issues and begin to develop recommendations
- **A series of virtual convenings** of targeted groups—judges, commissioners, attorneys, and other JPE stakeholders—to dive deeper on questions and key issues raised in the JPE Perspectives Survey and the in-person convening
- **This report, a JPE Perspectives Survey**, administered to judges in eight states, including Alaska, Colorado, Hawaii, Idaho, Massachusetts, New Mexico, Utah, and Virginia
- **Final recommendations and best practices** to improve JPE processes nationwide, informed by the preceding steps.

This JPE Perspectives Survey, detailing survey findings from judges across eight states, provides critical information to inform the final recommendations. Individual state reports have also been provided to each of the eight states to support their own improvement efforts.

Existing research on JPE largely focuses on critiques and potential improvements of evaluation programs. (For a summary of that research, please see Appendix A.) This study aims to fill a key gap in the existing research by offering the much-needed perspective of judges. The purpose of this survey was to gain a deeper understanding of judges' perspectives on the JPE process in their state today, what is working well, and challenges with the program. It also allowed us to gain a broader view of judges' perspectives across states. Gathering present perspectives of those most affected by and intimately familiar with JPE programs is critical to setting the foundation for potential improvements. If judges do not trust the processes through which they are evaluated, they are unlikely to modify their behavior in response to feedback produced by that process. It is crucial that judges trust and buy into the process. The sections that follow outline the JPE process and characteristics in different states, the methodological approach for the survey, and the survey results.

JPE PROGRAMS IN STUDY STATES

JPE programs vary considerably by state, both in purpose and scope, and this is true of the states involved in this study. For example, some programs conduct evaluations solely for the purpose of fostering judges' professional development and self-improvement, while others are designed to provide information about a judge's performance to elected officials or to the public. In addition, states vary in the tools and methods used to assess a judge's performance. The characteristics of study states are shown in Table 1. For more detail about each participating state's JPE programs, please see Appendix B.

Table 1: Participating State JPE Characteristics

	Commission	Attorney Surveys	Litigant Surveys	Juror Surveys	Court Staff Surveys	Opinion Review	Courtroom Observation	Objective Performance Standards	Judicial Discipline Records	Judge Self-Evaluation	Other Surveys
Alaska	✓	✓		✓	✓	✓			✓	✓	✓
Colorado	✓	✓	✓	✓	✓	✓	✓			✓	
Hawaii	✓	✓		✓							
Idaho		✓			✓						
Massachusetts		✓		✓	✓						
New Mexico	✓	✓	✓	✓	✓	✓	✓	✓		✓	
Utah	✓	✓		✓	✓		✓	✓	✓		
Virginia		✓		✓	✓	✓	✓			✓	

METHODOLOGY

The JPE Perspectives Survey asked judges about their court, the most important qualities for a judge, their experience with JPE, their opinions on JPE and its helpfulness, their concerns about JPE, and their suggestions for improving JPE. The survey instrument is shared in Appendix C.

In selecting states for participation, we looked for diversity across several dimensions, including JPE program characteristics, geographic location, and population characteristics. State selection was also conditioned upon a state’s willingness to participate. Ultimately, we distributed the JPE Perspectives Survey in eight states: Alaska, Colorado, Hawaii, Idaho, Massachusetts, New Mexico, Utah, and Virginia. Table 2 provides details about survey administration and response rates in each of the participating states.

Table 2: Participating State Survey Administration Dates and Response Rates

	Survey Administration Month(s)	Number of Judges to Whom Survey Was Distributed	Number of Judges Who Completed Survey	Response Rate
Alaska	December 2021	73	24	32.9%
Colorado	December 2021 to January 2022	358	113	31.6%
Hawaii	December 2021	76	25	32.9%
Idaho	December 2021	147	57	38.8%
Massachusetts	December 2021	360	24	6.7%
New Mexico	January 2022 to February 2022	135	42	31.1%
Utah	December 2021	193	95	49.2%
Virginia	December 2021	451	171	37.9%
Unknown ¹²			107	
TOTAL		1,793	658	36.7%

Substantial majorities of respondents judges identified as white (82.7%) and male (64.5%).¹³ While data about the demographics of state judiciaries is limited—including that of the participating states—these statistics generally align with the makeup of the state judiciaries participating in the study and state judiciaries generally.¹⁴

A majority of respondents (61.1%) worked in trial-level courts of general jurisdiction, while about one-quarter worked in limited jurisdiction trial courts (27.8%); the remainder worked in appellate or other types of courts. One-third worked in courts that serve primarily urban areas (33.7%), while slightly smaller proportions worked in courts that serve suburban/exurban areas (29.0%) or rural areas (27.1%). Only about 10% worked in courts that serve the entire state. A broad majority (84.7%) of respondents had experienced at least one JPE process in their time on the bench.¹⁵

As is true of all research, the results presented in this report should be considered in conjunction with the study’s limitations. While we have been diligent in working to minimize bias in the design and administration of the survey, we cannot be certain the extent to which non-response bias and other similar factors may have impacted our results. Relatedly, because data on the demographics of the population of state court judges in each of the participating states is not available, it is not clear the extent to which the demographics of the respondent group represent the demographic characteristics of the broader population of state court judges in

the study states. It is also true that large portions of our responses come from Colorado, Utah, and Virginia—so these perspectives may be disproportionately represented. Lastly, it should be noted that while this report focuses on findings aggregated across the eight study states, there are variations in judges’ perspectives from state to state.

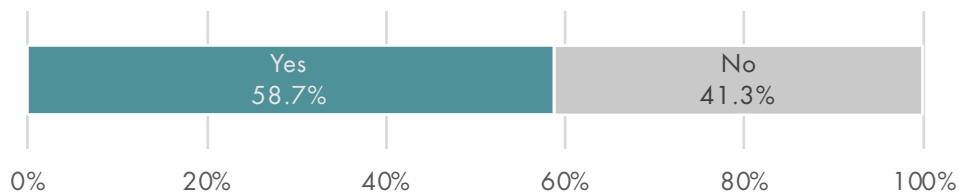
RESULTS

On the whole, survey responses reflect that judges largely have positive attitudes towards JPE in their states. For example, majorities of judges across study states reported that:

- They are satisfied with their state’s JPE process;
- The JPE process is beneficial to their professional development;
- The JPE process assesses their performance fairly; and
- Their final evaluation reports provide an accurate assessment of their performance.

However, despite these generally positive attitudes, a majority (58.7%) of respondents indicated that they had concerns about JPE in their state.

Figure 1: Do you have specific concerns about the evaluation process?¹⁶



This finding is further borne out in the narrative responses, where judges’ comments were largely negative. Indeed, sentiment analysis of the responses to the open-ended survey questions reveals that a larger proportion of comments contained negative sentiment than did positive sentiment by a factor of more than two—that is, a total of 435 of the open-ended survey response were coded as expressing negative sentiment, compared with 198 that were coded as expressing positive sentiment.¹⁷ While it is wise not to give much weight to counts derived from qualitative data, the sentiment analysis underscores the fact that, while there are broad indications of judges’ satisfaction with JPE processes, there is ample room for improvement.

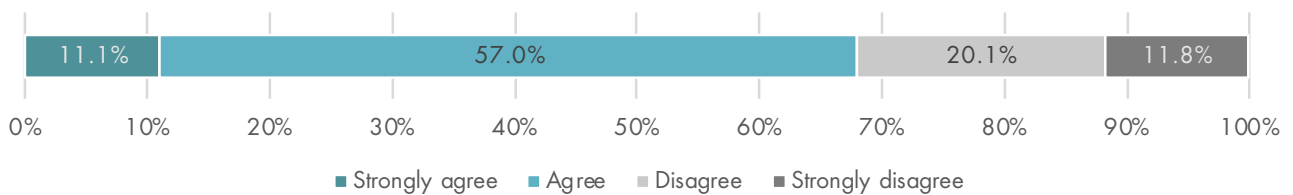
The sections that follow provide deeper insights into judicial perspectives on JPE—including perspectives on the JPE process, evaluation tools, evaluation results, commissions, and important qualities of a judge.

Perspectives on the JPE Process

JPE PROCESS: THE NUMBERS

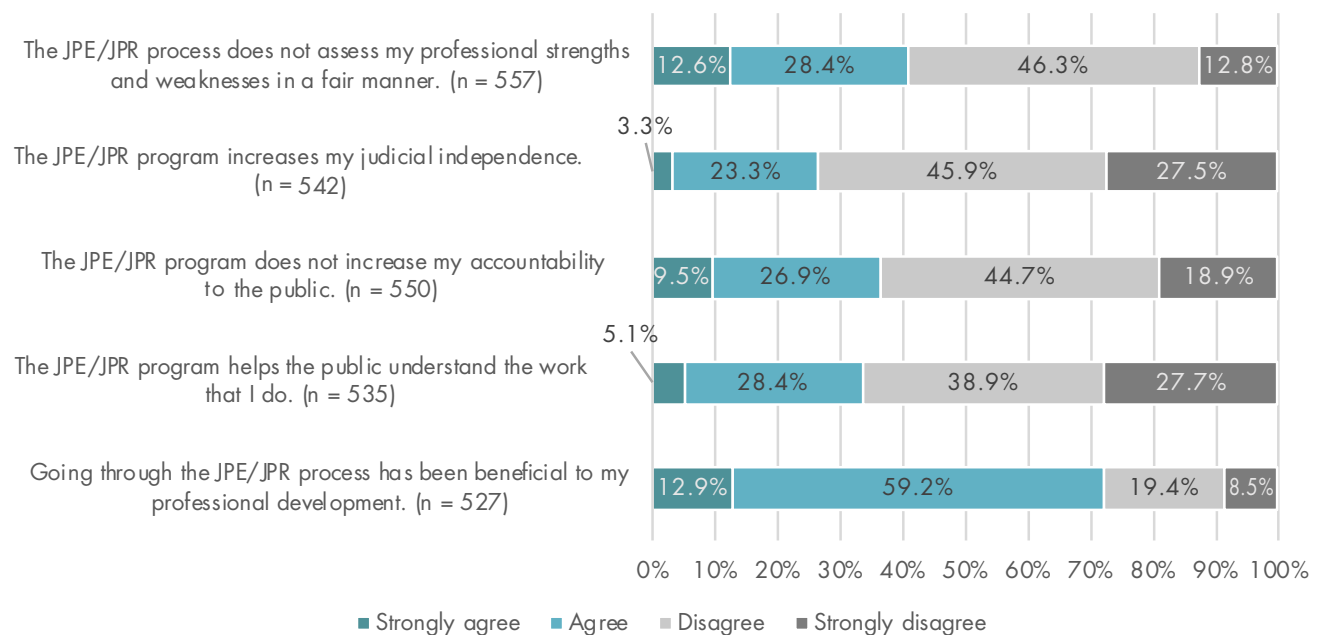
Across study states, just over two-thirds (68.1%) of judges reported that they were satisfied with the JPE process in their state, overall.

Figure 2: Overall, I am satisfied with the JPE/JPR process (n = 532)



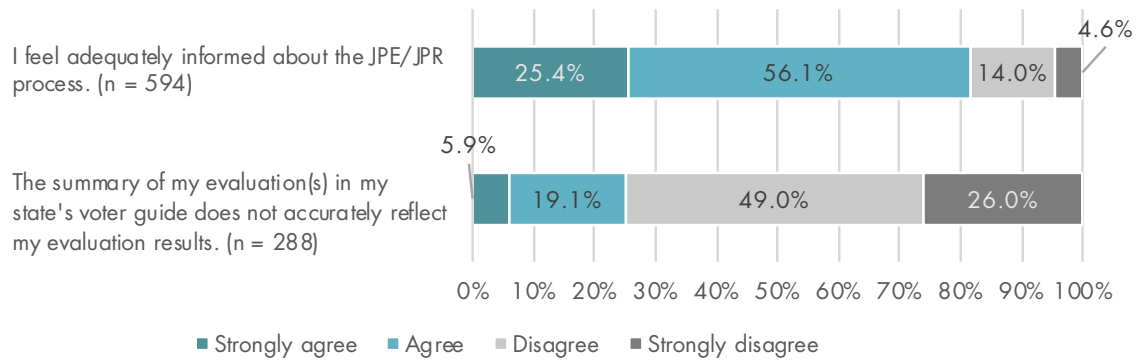
Digging more deeply, respondents recognized where JPE programs are creating positive outcomes and where such programs may be lacking.¹⁸ Specifically, majorities of respondents agreed that JPE has resulted in positive impacts with respect to serving as a fair assessment of judicial strengths and weaknesses (59.1%), increasing judicial accountability to the public (63.6%), and benefitting judicial professional development (72.1%). However, judges tended not to believe that JPE increases judicial independence (26.6%) or helps the public understand the work that judges do (33.5%).

Figure 3: Perspectives on the Impacts of JPE



With respect to communication about JPE, majorities of respondents felt adequately informed about the JPE process (81.5%) and believe that the evaluation summary in their state's voter guide accurately reflects JPE results (75.0%).

Figure 4: Perspectives on Communication about JPE



JPE PROCESS: THE NARRATIVES

Analysis of judges’ responses to open-ended survey questions illuminates how judges view the JPE process in terms of what is currently working, where improvements could be made, and specific suggestions from respondents.

WHAT IS WORKING	WHAT COULD BE IMPROVED	SUGGESTIONS FROM THE BENCH
Necessity and helpfulness of feedback	<ul style="list-style-type: none"> Fear and stress surrounding the process Gender and race-based bias in the evaluation process Inability for judges to respond to or contextualize comments Lack of public awareness about the process Negative impact on judicial independence 	<ul style="list-style-type: none"> Feedback from other judges Video review

WHAT IS WORKING

Necessity and Helpfulness of Feedback

Judges reflected that the feedback obtained through JPE was necessary for improvement and that it was generally helpful to receive such feedback. Judges expressed a desire to improve and get better at their jobs, and often viewed JPE evaluations as the only opportunity for a judge to receive invaluable and actionable insight. Different judges identified different evaluation tools and methods as most helpful. However, several judges noted that the process was especially helpful when a trusted mentor judge worked with them to implement improvements based on the feedback provided.

WHAT COULD BE IMPROVED

Fear and Stress Surrounding the Process

Many judges regard the JPE process as a source of stress and anxiety. Judges frequently described the process using terms like “dreaded,” “degrading,” “a source of consternation and alarm,” and “a significant negative impact on judges['] wellbeing.” For many reasons—such as issues with response rates, bias, and others as outlined elsewhere in this report—judges do not always trust the process to evaluate them fairly, and yet the stakes are very high if they receive poor evaluations.¹⁹ One respondent noted that, “When a judge’s career and livelihood are on the line in a retention election, such flaws in the evaluation process greatly exacerbate that stress.” The stress associated with the process leads judges to worry about losing their jobs and may cause them to make certain decisions out of fear of public repercussion or negative evaluations.

Gender and Race-Based Bias in Evaluations

For many judges, the pervasiveness of gender and racial bias undermines the integrity of the process. Several respondents noted that female judges and judges of color tend to experience the evaluation process differently due to implicit or explicit bias by attorneys, litigants, and commissioners. For example, one judge noted that she receives comments that she knows men on the bench would not receive, including remarks about her appearance and whether she is “acting nice enough.” Similarly, a judge of color noted that they have “on occasion received race-based commentary . . . that [has] been offensive and inappropriate.” When those comments are provided to the judge unfiltered, it can be “deflating.”

Inability for Judges to Respond to or Contextualize Comments

Judges reflected that when they are treated unfairly, take issue with a specific comment, or disagree with the results of an evaluation, there is little recourse. Several judges noted that there are not opportunities to raise legitimate concerns about the process or respond to negative feedback “without it reflecting poorly on them in the evaluation process.” One judge shared that they “had a specific, very negative comment in regard to [their] handling of a specific and unusual type of hearing in which [they had] never participated with no ability to respond to the comment.” This speaks not only to the accuracy of evaluations but also to fairness of the process for judges.

Lack of Public Awareness about the Process

Judges feel the process would be more effective if the public had greater awareness and education about how it worked. Judges expressed that the public generally does not understand the process, and many suggested that it would be helpful to provide more context for evaluation results.

Negative Impact on Judicial Independence

Some judges commented that the JPE process undermines or erodes judicial independence.²⁰ Rather than promoting self-improvement for judges, participants fear the current system leaves judges vulnerable to attacks that are motivated by case outcomes or politics rather than objective metrics of a judge’s performance, sometimes with “no form of quality control.” This fear of public backlash can incentivize judges to make rulings for the wrong reasons and deter well-qualified individuals from pursuing a career as a judge. Relatedly, judges from states where JPE data is provided to the legislative branch for reappointment purposes had concerns about the way legislators politicize that data and the resulting implications for the independence of the judicial branch.

SUGGESTIONS FROM THE BENCH

Feedback from Other Judges

Several judges suggested incorporating peer review from fellow judges on the bench. Other judges suggested that feedback and guidance from retired or mentor judges would be helpful given their experience. They felt that feedback from these colleagues would be most helpful, as other judges understand what the job entails and are unlikely to harbor bias or ulterior motives.

Video Review

One suggestion judges had to address the noted critiques and improve the overall JPE process was to incorporate video review as an evaluation tool, ideally by more experienced mentor judges. One judge suggested, “Now that we are conducting hearings on Zoom, those recordings could easily be accessed and provided to more experienced judges for quality control.”

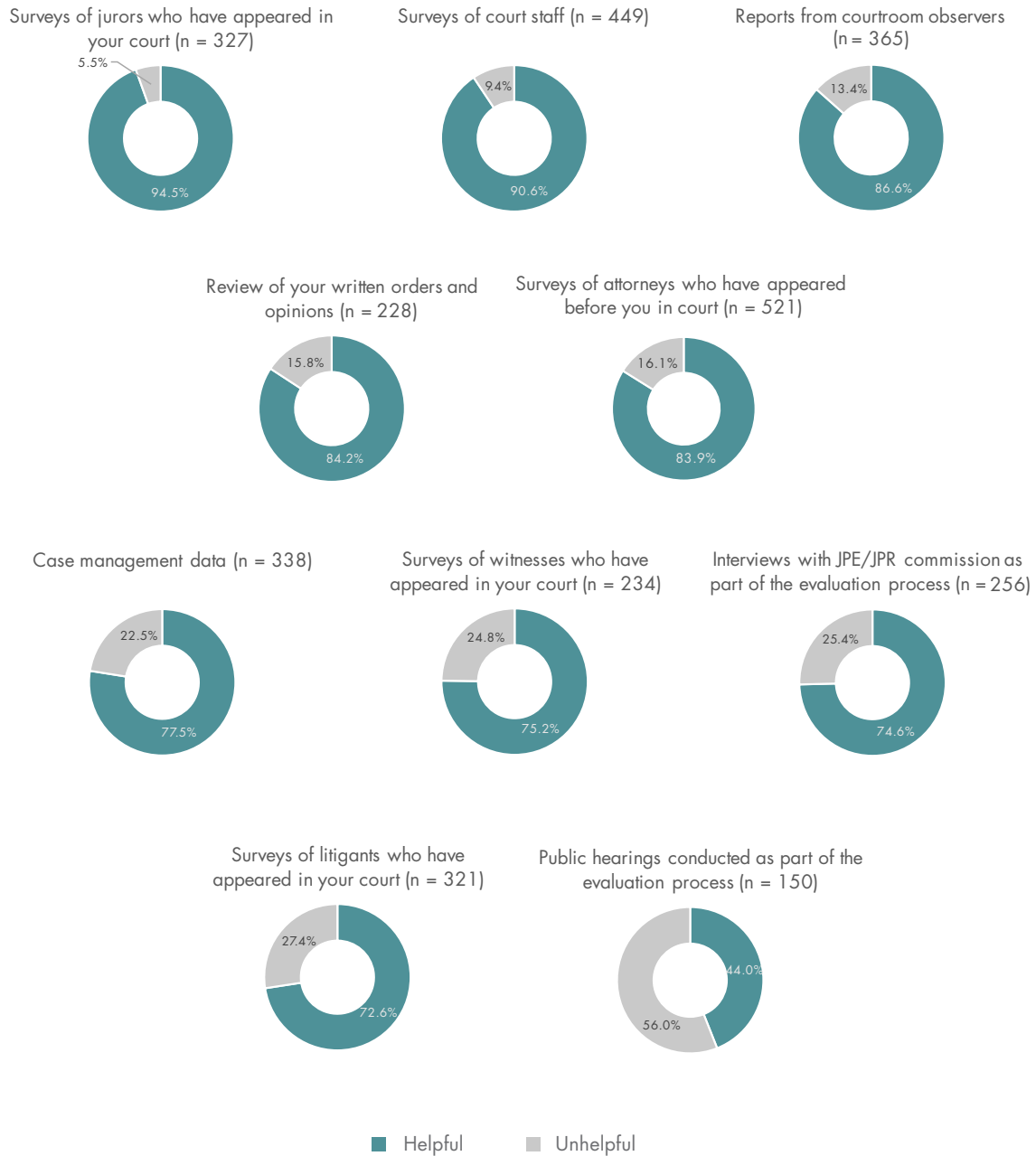
Perspectives on JPE Evaluation Tools

The JPE Perspectives Survey also asked judges to assess the helpfulness of each evaluation tool utilized in their state’s JPE process. Note that only those evaluation tools relevant to a given jurisdiction are reported (e.g., the figures only report responses about surveys of jurors in states where those are utilized as a part of JPE).

JPE EVALUATION TOOLS: THE NUMBERS

Substantial majorities of judges found each of the JPE evaluation tools to be helpful²¹—with the exception of public hearings, which less than half (44.0%) of judges reported were helpful. Among the most helpful evaluation tools were surveys of jurors (94.5%), surveys of court staff (90.6%), reports from courtroom observers (86.6%), review of written orders and opinions (84.2%), and surveys of attorneys (83.9%).

Figure 5: Perspectives on the Helpfulness of Evaluation Tools



*The “helpful” category combines responses that indicated the tool was both “very helpful” and “helpful.”
The “unhelpful” category combines responses that indicated the tool was both “unhelpful” and “very unhelpful.”*

JPE EVALUATION TOOLS: THE NARRATIVES

Judges' comments about JPE evaluation tools revolved around two specific tools: surveys and courtroom observations.

Surveys

The bulk of judges' comments on the evaluation tools utilized in the JPE process related to surveys—including surveys of attorneys, litigants, jurors, court staff, and others. These comments illuminate what judges view as working well and what they view as needing improvement, as well as suggestions for making improvements.

WHAT IS WORKING	WHAT COULD BE IMPROVED	SUGGESTIONS FROM THE BENCH
Written comments (in certain circumstances)	Low response rates Anonymity and vagueness of responses Weaponization of surveys Negative bias and outcome bias of respondents Identity and capacity of respondents	Screening and weighting of responses More constructive feedback

WHAT IS WORKING

Written Comments (in Certain Circumstances)

Positive responses related to surveys were minimal, but some respondents noted that written comments in surveys, and specifically comments from certain groups, were helpful to them. Although written comments tend to have a negative bias and are not always constructive, some judges found them useful in making modifications to the way they run their courts. More judges, it should be noted, felt that the comments are unhelpful, as the concerns about bias and personal vitriol detailed later outweigh their utility. Even judges who found comments helpful noted in the same response that there were significant limitations. Several judges commented that the feedback from lawyers is most helpful, though sentiments about lawyer feedback were certainly mixed. One judge summed up the complex views on this topic by saying, “[S]pecific comments are most helpful, but this is tempered by those with an axe to grind[.]”

WHAT COULD BE IMPROVED

Low Response Rates

Low survey response rates are a top concern across jurisdictions. Several judges commented that low response rates lead to unreliable results, making it hard to assess patterns, and “do[] not give a good picture of the performance of each judge.” Often, evaluators draw conclusions about a judge’s performance based on survey responses from a limited number of participants. Small sample sizes tend to mean that responses from individual attorneys can have an outsized impact on a judge’s evaluation results. Negative comments are weighted more heavily because, in the view of many judges, survey respondents tend to be a “self-selecting sample” of “those who were most unhappy with the result in their case.”

Anonymity and Vagueness of Responses

Judges frequently commented on the challenges raised by the anonymity of survey responses. While anonymity plays an important role in encouraging survey participation and protecting participants from retaliation for negative reviews, judges often feel the anonymity makes it difficult to understand the context behind a comment and discern which critiques have merit. This in turn undermines the helpfulness of the feedback. It can also embolden disgruntled attorneys and litigants to make “venomous,” “disingenuous,” or “bullying” comments. One respondent described interpreting anonymous comments as “chasing ghosts.” The anonymity also means there is no verification process to ensure that the commenter actually has personal experience with the judge being evaluated.

Weaponization of Surveys

Many judges expressed that anonymous and unvetted surveys are susceptible to being weaponized by attorneys to harm a judge’s reputation through “unfettered and often baseless attacks.” One judge shared that they have “witnessed lawyers conspiring to submit similar comments about judges, in an effort (successfully) to have the judge receive a ‘do not retain’ recommendation.” This is exacerbated by the fact that there are no checks on comments, such as to ensure the respondent has actually appeared before the judge. The weaponization of surveys “discourages well qualified lawyers from considering a career on the bench” and “creates judges who make rulings . . . based on concerns over the judicial evaluation process and whether their careers will be destroyed because a lawyer[] doesn’t like them.”

Negative Bias and Outcome Bias of Respondents

Judges from several different states shared the view that people are most likely to respond to surveys when they are unhappy, and that those who have positive experiences with a judge are less likely to respond. Respondents expressed concern that this skews survey results, as it gives negative comments more weight and offers “disgruntled attorneys a larger chair at the table.” Relatedly, judges observed that individuals tend to use surveys to “evaluate the outcome of a case” when they receive an adverse ruling rather than evaluating the judge’s performance. One judge attributed this to misunderstanding of the process: “[L]itigants do not understand that they are evaluating the judge . . . they evaluate the outcome of the case.” Other judges chalked this up to respondents evaluating judges “based upon personal feelings” or because they have a “proverbial axe to grind.”

SUGGESTIONS FROM THE BENCH

Screening and Weighting of Responses

Judges' suggestions related to improving surveys centered on better filtering, for example by weighting, removing, or qualifying certain responses or by fact checking them. Several judges suggested that responses containing hateful or personal attacks should either be disqualified or screened out before results are shared with the judge. Speaking to hurtful attacks, one judge said, "I don't need to see those comments." In addition, judges felt there should be a way to qualify comments that are clearly motivated by an adverse ruling in a case and not related to a judge's actual performance. Some judges also felt that commissions could do a better job weighting or rating comments from different categories of respondents. For example, comments from an attorney who appeared before a judge many times could receive greater weight than comments from an attorney who appeared before that judge only once. Similarly, a couple of judges urged quality control mechanisms to ensure attorneys responding to surveys have actually appeared before the judge.

More Constructive Feedback

It was clear throughout their responses that judges want to improve. Many judges noted that the surveys could do a better job of encouraging constructive feedback, such as by asking more specific questions about what a judge could do to improve rather than open-ended questions. This would allow judges to better implement feedback from evaluations.

Courtroom Observations

Judges also shared many thoughts related to what is working well, what could be improved, and suggestions for making improvements to how courtroom observations are used as evaluation tools in JPE processes.

WHAT IS WORKING	WHAT COULD BE IMPROVED	SUGGESTIONS FROM THE BENCH
Observers	Duration and frequency of observations Training of observers	Increased involvement of judges in observations

WHAT IS WORKING

Observers

The primary positive feedback on courtroom observations was about the observers themselves, though the identity of observers varies depending on the state (e.g., commissioners, other judges, etc.). Judges noted that while courtroom observers are "generally helpful" because they are "unbiased," the level of helpfulness is "highly dependent on their training and experience." One respondent noted that "[t]he observers are not

intrusive and seem well-intentioned.” Another stated that review from a senior judge observer was most helpful because “[t]heir assessments . . . come from a desire to assist the judge to improve and are not clouded by whether they won or lost the case.”

WHAT COULD BE IMPROVED

Duration and Frequency of Observations

Several judges took issue with the short period of time during which they are observed, noting that observing for just a few hours or even a couple of days does not provide a full picture of how a judge runs their courtroom. One judge commented that “relying on such limited knowledge to make such an important recommendation is a disservice to the community and the judicial officer who is being evaluated.” Other judges commented that they are not observed at all, and that they wished to be. Despite the critiques, these comments indicate that judges find courtroom observations useful as an evaluative tool and generally want to see them used more.

Training of Observers

While judges generally found comments from courtroom observers helpful because they tend to be unbiased, they expressed a desire for better training for observers, specifically about the judge’s role, the demands of their specific docket, and court processes. One judge described it as “frustrating when the observers don’t have any courtroom experience.” Others reflected that even when observers do receive training, it may be insufficient for them to effectively evaluate judges. For example, another judge said, “I am not confident that the training that court evaluators receive is sufficient to achieve a fair, relevant report. I have received comments on past reviews that make it clear to me that the evaluators don’t understand our court processes well enough to comment on how well I am performing them.”

SUGGESTIONS FROM THE BENCH

Increased Involvement of Judges in Observations

To address the noted critiques, a few respondents suggested it would be helpful to involve judges more in the courtroom observation process. One judge said, “I believe I would learn from watching other judges and providing feedback while reflecting on my own practice and feedback received.”

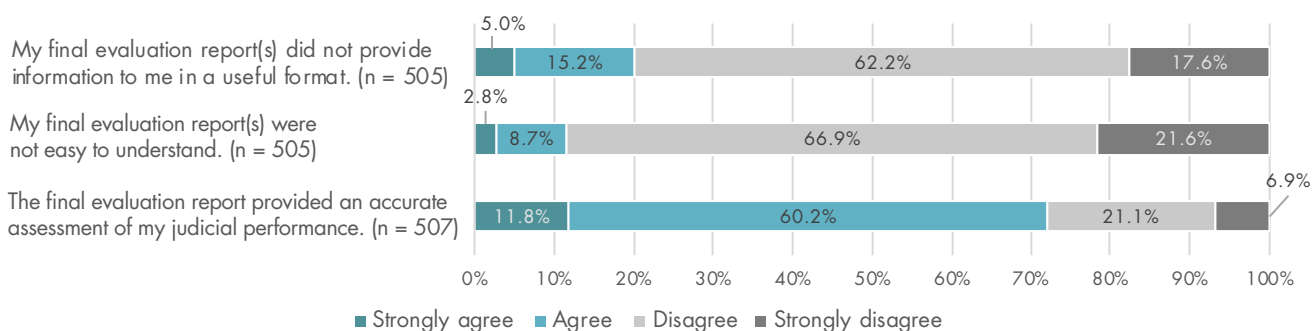
Perspectives on JPE Evaluation Results

The JPE Perspectives Survey asked judges to reflect on the final evaluation reports provided at the end of the JPE process in terms of the format and content of those reports, as well as the feedback contained within them.

JPE EVALUATION RESULTS: THE NUMBERS

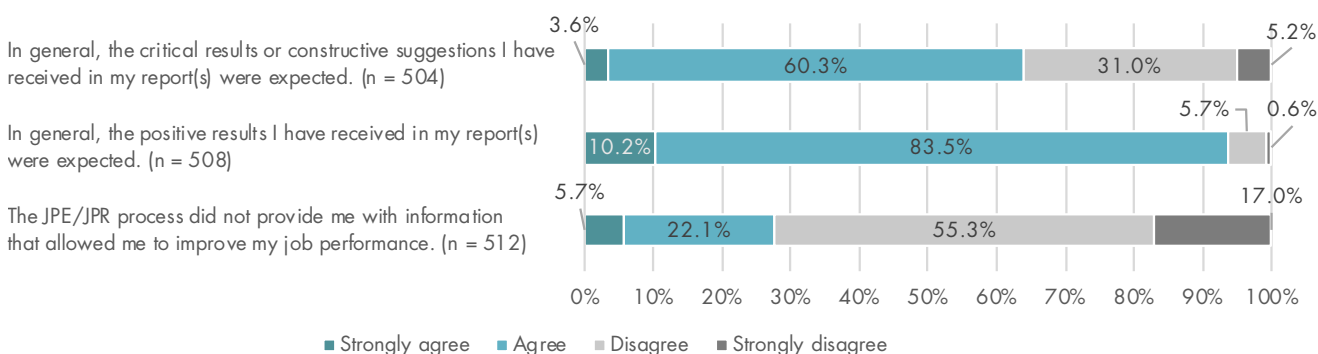
Judges largely had positive views about the final evaluation report provided to them at the conclusion of the JPE process. Substantial majorities of respondents thought that their final report provided information in a useful format (79.8%), was easy to understand (88.5%), and provided an accurate assessment of their performance (72.0%).

Figure 6: Perspectives on the Final Evaluation Report



With respect to the feedback and information received in the final evaluation report, considerable majorities of judges were unsurprised by both their critical and positive results (63.9% and 93.7%, respectively). Nearly three-quarters of judges (72.3%) thought that the JPE process provides information that allows them to improve their job performance.

Figure 7: Perspectives on Feedback Received



JPE EVALUATION RESULTS: THE NARRATIVES

Respondent judges provided deeper insights into what is working, what could be improved, and suggestions for improving how feedback and final JPE results are shared with judges.

WHAT IS WORKING	WHAT COULD BE IMPROVED	SUGGESTIONS FROM THE BENCH
Helpfulness of feedback and evaluation reports to judges	Consistency of evaluations	Review of written orders and opinions
Helpfulness of public reports	JPE as popularity or personality contest	Inclusion of quantifiable metrics to supplement subjective feedback
	Lack of recourse for judges	More specific feedback
	Lack of accountability for evaluators	Private evaluations
	Complexity of public reports	No recommendation on retention
	Inconsistency of public reports	Better resources for implementing feedback
	Recommendations for growth	Ongoing tracking of improvements

WHAT IS WORKING

The Helpfulness of Feedback and Evaluation Reports to Judges

As noted many times throughout survey responses, judges want feedback on their performance. A couple of judges commented that although there are flaws with JPE and a judge's performance is difficult to evaluate, they generally felt the feedback was helpful and were supportive of an evaluation process that could be used to improve their performance.

The Helpfulness of Public Reports

Regarding the narrative provided to the public, a handful of judges acknowledged that it can provide helpful information to the public, with one stating "the report written by the commission is very helpful in all aspects."²²

WHAT COULD BE IMPROVED

Consistency of Evaluations

Different judges experience the JPE process differently, such as based on their gender or race or based on who is evaluating them. This leads to a general sense of unfairness, especially for women and judges of color, who participants felt "are not always treated fairly."

JPE as Popularity or Personality Contest

Several judges commented that the JPE process feels like a “popularity contest” or a “personality contest” in that judges who are likeable receive better evaluations. One judge said, “[T]he evaluations are a popularity contest. People either like you or they don’t.” Even judges who receive generally positive evaluations felt this does not yield constructive feedback helpful to job performance.

Lack of Recourse for Judges

A couple of judges felt the feedback was “one-sided” with little opportunity for recourse when a judge is treated unfairly. One judge remarked, “Results can be the equivalent of cyber-bullying and judges have virtually no recourse or due process rights.”

Lack of Accountability for Evaluators

Several judges commented on the lack of oversight and accountability when it comes to what evaluators include in the public reports or voter guides. One judge noted, “My experience has been that most of the Commissions will present fair, thoughtful and constructive comments for the voter guides. However, when a Commission exceeds their role, a judge needs to have some means outside of the Commission to obtain a fair review of the Commission’s comments.”

Complexity of Public Reports

A few of the judges responded that the data in public reports was helpful but is presented in a way that is difficult to interpret, such that “the public doesn’t understand these reports.” As one judge remarked, they “could be presented in a format that is easier to digest” for the public.

Inconsistency of Public Reports

As noted previously with respect to JPE results, a couple of judges commented that the narrative reports or voter guides provided to the public can be inconsistent depending on a judge’s identity, the nature of their court, and the identity of the evaluator writing it. One judge remarked, “The comments in the voter guide largely depends on who writes the narrative. Some take out great quotes and provide glowing remarks while others are very factual and dry leaving a poor impression in comparison.”

Recommendations for Growth

Several judges in different jurisdictions responded that they would like more actionable feedback and better opportunities to work with others to address concerns about their performance. This is consistent with the theme that came up many times throughout the judges’ remarks: they genuinely want to improve. However, the JPE process does not always facilitate this improvement successfully. (It should be noted that sentiments about this seem to be dependent on jurisdiction, as the judges who made these critiques were from different states than the judges who had positive comments about the review of feedback in their JPE programs.)

SUGGESTIONS FROM THE BENCH

Review of Written Orders and Opinions

Several judges noted that objective review of written orders and opinions, specifically by a more experienced judge, would be helpful. Judges also suggested assessing additional characteristics of decisions such as whether those decisions were well-reasoned, whether their sentences were consistent for similar crimes, how many appeals were filed, and what the outcome was of those appeals.

Inclusion of Quantifiable Metrics to Supplement Subjective Feedback

A few judges suggested adding in more objective metrics to the evaluation process to supplement attorney opinions. For example, commissions could review a judge's case management statistics related to metrics like timeliness of resolution, length of trials, and number of cases.

More Specific Feedback

As was a theme in many of the judge's responses, they are eager for more specific and constructive feedback that includes "examples of performance issues" and suggestions for improvement.

The Role of Private Feedback and Mentorship

A few judges commented that feedback would be more helpful if it were private and accompanied by mentorship rather than made available to the public. One judge stated that private feedback would "be more productive to helping improve performance."

No Recommendations on Retention

Several judges argued that the JPE results are "not scientific" and "not probative of anything" so should not ultimately impact a judge's position on the bench. Rather, they should be used solely for judicial self-improvement.

Better Resources for Implementing Feedback

Many judges from various jurisdictions suggested more robust resources and support for improving their performance in light of feedback from the evaluations. One said, "The process does not allow or encourage ongoing feedback or constructive criticism that would allow me to make changes." There were many specific suggestions made about how to do this, including the use of mentor judges to help address concerns raised during review, more one-on-one interactions with evaluators centered around strategies for improvement, and practical training for judges on how to improve the court experience. One judge summed up this sentiment by saying, "If a judge is graded negatively on a particular score, it should be a requirement that the evaluator provide guidance as to how the judge could improve as to that particular score."

Tracking of Ongoing Improvements

Relatedly, several judges urged that the JPE evaluations should be an ongoing process that involves feedback combined with support and the opportunity to improve performance. One judge said, "Judges need to be rated on actual empirical improvement, not aspirational goals[.]"

Perspectives on Commissions

The JPE Perspectives Survey did not specifically ask for feedback on JPE commissions or commissioners. However, judges frequently provided both positive and critical narrative feedback on JPE commissions.

WHAT IS WORKING	WHAT COULD BE IMPROVED	SUGGESTIONS FROM THE BENCH
Commission membership and leadership	Bias and political agendas Interpretation of results	More constructive feedback Judges as evaluators

WHAT IS WORKING

Commission Membership and Leadership

Though reviews were certainly mixed, some judges identified the commission members and leadership as a positive and an important factor in the success and validity of a JPE program. Where they had positive things to say about the commissions, they noted their professionalism, diversity, and balance. When JPE leadership is perceived as competent, reliable, and harboring good intentions, judges are better able to trust that the process is fair. Of course, the converse is also true. When JPE leadership is perceived as biased, unreliable, and lacking knowledge of the process, judges are skeptical of the evaluations and believe them to be unfair.

WHAT COULD BE IMPROVED

Bias and Political Agendas

Despite these positive sentiments, judges shared many critiques of commissions, particularly related to bias, political agenda, and inaccurate interpretation of results. For example, several respondents from multiple states noted that commissions sometimes lack proper training or monitoring when it comes to bias. In addition, they felt that some commissioners appear to have an agenda, either in general or when it comes to evaluating a particular judge, which “is not consistent with a fair and impartial review of a judge’s performance.”

Interpretation of Results

Furthermore, several judges responded that commissioners present all evaluation results to the judge and the public as truth, even when it is inaccurate, baseless, or a mischaracterization. One judge shared, “My evaluation included inaccurate information that . . . commissioners failed to correct despite my bringing it to their attention. They were more focused on giving me negative reviews than in evaluating me fairly.” These dynamics make it difficult for judges to trust that the process is fair.

SUGGESTIONS FROM THE BENCH

More Constructive Feedback

As noted before, judges genuinely want to improve and are eager for constructive critiques, even if they are uncomfortable to hear. One judge said, “The judicial evaluation process serves a useful purpose when the evaluators provide a constructive review of the judge’s performance. For instance, ‘the judge needs to make better eye contact with the defendant in court.’ I appreciate this constructive criticism even when it stings a little if I know it to be an honest evaluation.” Commissions should filter out unhelpful or overly personal criticism and focus on specific suggestions for how a judge can make necessary improvements.

Judges as Evaluators

In order to effectuate this, several respondents noted that it would be helpful to incorporate peer or mentor judges into the evaluation process. “Who better to judge a judge?” one respondent said. Another said, “We all have room to improve and with the guidance of an experienced judge, improvement could happen.”

Most Important Qualities in a Judge

In addition to asking about various aspects of JPE, the JPE Perspectives Survey asked judges—in an open-ended question—to indicate the top three most important skills, abilities, or qualities that make a good judge. Coding and analysis of these responses demonstrates that judges recognize the importance of possessing a combination of legal expertise, professional skills, and interpersonal abilities. Table 3 presents the top 20 coded skills, abilities, and qualities judges identified as most critical for success.

Table 3: Top 20 Skills, Abilities, and Qualities for a Good Judge

Name	Coded References
Patience	300
Knowledge	292
Neutrality and fairness	197
Compassion	129
Intelligence	88
Demeanor	87
Temperament	81
Civility and respect	74
Communication	56
Integrity and trustworthiness	56
Organization	52
Humility	48
Work ethic	46
Diligence	40
Decisiveness	39
Experience	35
Legal analysis and ability	33

DISCUSSION AND CONCLUSION

The research in this report builds upon a 2008 IAALS study designed to measure the effectiveness of Colorado's JPE program, *The Bench Speaks on Judicial Performance Evaluation: A Survey of Colorado Judges*.²³ The results of that survey revealed that Colorado judges were “quite comfortable with the concept of judicial performance evaluation and measurement of job performance based on process-oriented criteria” as well as the collection of a wide range of data to evaluate their job performance, so long as those evaluations are “conducted fairly and reasonably.”²⁴ In that survey, over 85% of trial judges and 50% of appellate judges answered that JPE had been either “significantly beneficial” or “somewhat beneficial” to their professional development. In addition, most judges did not feel JPE decreased their judicial independence.²⁵ Judges did express concerns about the way evaluations were conducted and circulated. First, judges felt that some evaluations were based on unreliable survey data. Second, judges suggested that the public needs to be made more aware of evaluation results and how to access them. These findings are generally consistent with the survey results contained in this report conducted almost 15 years later.

Fast forward to today and the complex challenges facing our world, our courts, and our judges. Like the 2008 Colorado study, the findings from this current JPE Perspectives Survey present a seeming duality: on one hand, the numbers paint a picture of a judiciary largely satisfied with JPE processes but, on the other hand, judges' comments and survey responses point to many serious, legitimate concerns.

Areas for Continuous Improvement

OVERALL PROCESS

Judges are largely satisfied with the overall process, and they generally see the need for this feedback and think the JPE process is effective in serving as a fair assessment of judicial performance, increasing judicial accountability and promoting judicial professional development. Judges express concerns related to fear and stress surrounding the process, gender and racial bias impacting results, inability of judges to provide responses or context to comments, lack of public awareness about JPE, and negative effects on their judicial independence.

EVALUATION TOOLS

Judges, in general, indicated that the evaluation tools utilized during the JPE process are helpful. With respect to surveys, many thought that written comments can be helpful—but the circumstances under which they are considered helpful vary. By and large, though, judges' comments related to challenges associated with the surveys, including low response rates, anonymity and vagueness of responses, perceptions that surveys can be weaponized, bias in responses, and which groups are surveyed. Many judges also commented on courtroom observations; observers are often well regarded, but several judges raised concerns about the duration and frequency of observations, as well as the training of observers.

EVALUATION RESULTS

In terms of the evaluation results, judges generally had positive views of the format, understandability, and accuracy of their final evaluation reports they receive from commissions. They also generally found that the results aligned with their expectations and were helpful information both for themselves and the public. According to respondents, areas of improvement related to the evaluation results include consistency of evaluations, the perception that JPE is a popularity or personality contest, lack of recourse for judges who feel a result is unfair, and lack of accountability for evaluators, among others.

JPE COMMISSIONS

Judges also shared perspectives on JPE commissions. They frequently shared positive views about the commission members and leadership and negative comments about commissions generally related to perceptions of bias and political agendas, as well as how commissions interpret evaluation results. JPE commissions and judges are aligned in their goals; both want to gain tools for judges to do their jobs better and to improve public trust and confidence in our court systems. It is critical that commissions work alongside judges and consider their feedback in order to make JPE work better for everyone.

The results from this JPE Perspectives Survey provide critical insights from the judges who engage with JPE processes, and they provide a roadmap for states to improve their JPE programs. Judges' perspectives on JPE present a vision of a fairer, more effective, and more useful evaluation process.

This study, paired with the perspectives of JPE stakeholders that IAALS has gathered from across the country, will inform the next JPE 2.0 report containing updated recommendations and best practices to guide JPE programs into the future. These recommendations will equip states to improve and optimize their JPE programs, and make them work better for all involved—the judges, the program administrators, and the public.

ENDNOTES

- 1 For simplicity, these programs are referred to as “JPE” throughout the report except where specific survey questions included “JPR.”
- 2 The following states currently have some version of a JPE program: Alaska, Arizona, Colorado, Connecticut, Hawaii, Idaho, Illinois, Massachusetts, Missouri, New Hampshire, New Jersey, New Mexico, Rhode Island, Utah, Vermont, and Virginia.
- 3 JORDAN M. SINGER, INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., JUDICIAL PERFORMANCE EVALUATION IN THE STATES: THE IAALS JPE 2.0 PRE-CONVENING WHITE PAPER 1-2 (2022).
- 4 *Id.* at 2.
- 5 David C. Brody, *The Use of Judicial Performance Evaluation to Enhance Judicial Accountability, Judicial Independence, and Public Trust*, 86 Denv. U. L. Rev. 115, 121-27 (2008).
- 6 *Id.* at 156-57.
- 7 LOGAN CORNETT & NATALIE ANNE KNOWLTON, INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., PUBLIC PERSPECTIVES ON TRUST AND CONFIDENCE IN THE COURTS 3 (2020).
- 8 SINGER, *supra* note 3, at 5.
- 9 For example, a study by the National Center for State Courts found that in 76% of civil cases, at least one party was self-represented. See NAT’L CTR. FOR STATE COURTS, THE LANDSCAPE OF CIVIL LITIGATION IN STATE COURTS iv (2015), <https://www.ncsc.org/~media/Files/PDF/Research/CivilJusticeReport-2015.ashx>.
- 10 INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., The O’Connor Judicial Selection Plan (2014), https://iaals.du.edu/sites/default/files/documents/publications/oconnor_plan.pdf.
- 11 See generally *id.*; NATALIE KNOWLTON & MALIA REDDICK, INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., LEVELING THE PLAYING FIELD: GENDER, ETHNICITY, AND JUDICIAL PERFORMANCE EVALUATION (2012), https://iaals.du.edu/sites/default/files/documents/publications/iaals_level_the_playing_field.pdf; REBECCA LOVE KOURLIS & MALIA REDDICK, INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., TRANSPARENT COURTHOUSE REVISITED: AN UPDATED BLUEPRINT FOR JUDICIAL PERFORMANCE EVALUATION (2016), https://iaals.du.edu/sites/default/files/documents/publications/transparent_courthouse_revisited.pdf; REBECCA LOVE KOURLIS, MALIA REDDICK, & NATALIE ANNE KNOWLTON, INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., RECOMMENDED TOOLS FOR EVALUATING APPELLATE JUDGES (2013), https://iaals.du.edu/sites/default/files/documents/publications/recommended_tools_for_evaluating_appellate_judges.pdf; Rebecca Love Kourlis & Jordan M. Singer, *Using Judicial Performance Evaluations to Promote Judicial Accountability*, 90 *Judicature* 201, 206 (2007).
- 12 There were 107 responses from individuals who did not indicate which state they were from but who completed the survey distributed to Alaska, Colorado, Hawaii, Massachusetts, Utah, and Virginia.
- 13 See Appendix C for detailed demographic breakdowns.
- 14 See Tracey E. George & Albert H. Yoon, *Measuring Justice in State Courts: The Demographics of the State Judiciary*, 70 *Vand. L. Rev.* 1887 (2017).
- 15 See Appendix C for detailed breakdowns of respondent courts and respondents’ experience with the JPE process.
- 16 Actual question text: “When you think about JPE/JPR in your state as a whole, do you have specific concerns about the evaluation process?”
- 17 In our sentiment analysis, each open-ended response could be coded as having positive, neutral, or negative sentiment. A single response could receive more than one of these three codes (for instance, a response could be coded for both positive and negative sentiment), but would not be coded multiple times for the same sentiment. These calculations exclude coded responses to a question that explicitly asked respondents to detail their concerns (because responses to this question are inherently negative).

- 18 Note that we worded survey statements both positively and negatively to help minimize bias in the survey instrument.
- 19 It should be noted, however, that 59.1% of judges said they either disagreed or strongly disagreed with the statement, “The JPE/JPR process does not assess my professional strengths and weaknesses in a fair manner.” See Figure 3.
- 20 In fact, 73.4% of judges disagreed or strongly disagreed with the statement, “The JPE/JPR program increases my judicial independence.” See Figure 3.
- 21 Respondents were asked to rate helpfulness on a four-point scale: very helpful, helpful, unhelpful, or very unhelpful. In Figure 4, responses have been collapsed into two categories: helpful and unhelpful. For a full breakdown of helpfulness ratings of each of the evaluation tools, see Appendix E.
- 22 In fact, 66.5% of participants agreed or strongly agreed with the statement that, “The JPE/JPR program helps the public understand the work that I do.” See Figure 3.
- 23 INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., THE BENCH SPEAKS ON JUDICIAL PERFORMANCE EVALUATION: A SURVEY OF COLORADO JUDGES (2008), https://iaals.du.edu/sites/default/files/documents/publications/bench_speaks_on_jpe2008.pdf.
- 24 *Id.*
- 25 About 41% of appellate judges and 44% of trial judges said that the JPE had no effect on their judicial independence. Another 24% of appellate judges and 29% of trial judges felt JPE *increases* their judicial independence.





Institute for the Advancement of the American Legal System
University of Denver
John Moye Hall, 2060 South Gaylord Way
Denver, Colorado 80208

303.871.6600 iaals.du.edu @iaals