



RECOMMENDATIONS FOR MODERNIZING 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

RECOMMENDATIONS FOR MODERNIZING JUDICIAL PERFORMANCE EVALUATION

Danielle Kalil

Director of Civil Justice and the Judiciary

May 2025

For reprint permission, please contact IAALS.
Copyright © 2025 IAALS, the Institute for the
Advancement of the American Legal System.
All rights reserved.



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

How can we help? Consulting Inquiries: iaals@du.edu

TABLE OF CONTENTS

Message to Readers	1
Introduction	2
Challenges and Opportunities Facing JPE Programs	4
Guiding Principles	8
Key Considerations.....	9
Recommendations for Modernizing JPE Programs	10
Recommendation 1: Foster trust and confidence in the JPE process.	11
Recommendation 2: Modernize the evaluation criteria to account for the changing roles and responsibilities of the judiciary	14
Recommendation 3: Implement measures to improve information and reduce errors in evaluations	19
Recommendation 4: Build a culture of and embrace judicial performance improvement as a key aspect of JPE	26
Recommendation 5: Display institutional commitment to JPE	27
Recommendation 6: Ensure and improve transparency for intended audiences	28
Call to Action: What Comes Next?	30
Conclusion.....	32

ACKNOWLEDGEMENTS

We thank the current and former members of the IAALS JPE 2.0 Task Force for their contributions and guidance in crafting these recommendations.

- ▶ **Barbara Arnold**, Former Program Manager, New Mexico Judicial Performance Evaluation Commission
- ▶ **Susanne DiPietro**, Executive Director, Alaska Judicial Council
- ▶ **Farah Fite**, Former Communications Director, The Missouri Bar and Staff Liaison, Missouri Judicial Performance Review Committee
- ▶ **Michael Oki**, Research Statistician, Hawai'i State Judiciary
- ▶ **Jordan Singer**, Professor of Law, New England Law | Boston
- ▶ **Kent Wagner**, Executive Director, Colorado Office of Judicial Performance Evaluation
- ▶ **Jennifer Yim**, Former Executive Director, Utah Judicial Performance Evaluation Commission

We also thank our additional partners who—along with Task Force members and state program administrators—helped develop, distribute, and review the JPE Perspectives Survey, the eight-state survey of judges that informed these recommendations:

- ▶ **Adam Cohen**, Planning and Program Evaluation Division, Hawai'i State Judiciary
- ▶ **Anna Rachel Dray-Siegel**, Assistant Legal Counsel, Massachusetts Supreme Judicial Court
- ▶ **Mateo Page**, Former Director, Court Services Division, New Mexico Administrative Office of the Courts
- ▶ **Andrea Powers**, Human Resources Director, Idaho Judicial Branch
- ▶ **Mary-Margaret Pingree**, Executive Director, Utah Judicial Performance Evaluation Commission
- ▶ **Kristi Wright**, Program Director, Virginia Judicial Performance Evaluation Program

We also thank current and former members of the IAALS JPE Working Group who have provided valuable on-the-ground insights that have informed this project:

- ▶ **Darrin Browder, Blanca Moreno Calles, Chelsea Canon, Dale Doerhoff, Hannah Kiddoo Frevert, Av Harris, Mona Hochberg, Mateo Page, and Gary Toohey**

In addition, we thank everyone who participated in our May 2022 convening and the series of 2023 JPE 2.0 virtual convenings for their insights and input on JPE programs, areas for reform, and recommendations.

Finally, we are grateful to **Russell Wheeler** and **Judge Adam Espinosa**, who provided valuable feedback on these recommendations.

MESSAGE TO READERS

IAALS is excited to share these recommendations to guide the future of judicial performance evaluation in the United States. They are the culmination of extensive research, collaboration with JPE administrators and judges across the country, and consideration of evidence-based best practices to modernize JPE. We believe the recommendations outlined here provide a roadmap to support the judiciary's position as a trusted institution. But before we get to the *what*, we want to share the *why* behind our many years of work on JPE.

JPE is about so much more than an evaluation. Throughout these recommendations, you will see themes emerge with applicability far beyond JPE: judicial independence, transparency and accountability to the public, public trust and confidence in our institutions, civility in public discourse, public civic engagement, professional responsibility in the legal profession, the role and reliability of data, and the role of technology.

These broad themes demonstrate that the recommendations in this report can have impact and applicability beyond the context of formal JPE programs. While JPE programs offer a structured approach to evaluating judicial performance, the principles underlying these recommendations are relevant across all of our courts and to our democratic institutions more broadly. They highlight the critical role of judges, the public's perception of the judiciary, the importance of investing in ongoing professional development for judges, and the need to improve public trust and confidence in courts.

To build public trust, we must create a justice system that is truly trustworthy. The fair, impartial, and efficient functioning of our justice system has never been more important. We must put everyday people at the system's center, and we must ensure that legal and court services are understandable, accessible, and affordable for everyone. Where the system is not meeting these goals, it undermines public trust and confidence of individuals in their democratic institutions, and it enables attacks at a system-wide level that undermine the justice system as a whole. We need a strong, independent justice system that meets the modern needs of people and society, now more than ever. Putting these principles into practice, whether in a state with JPE or without, will help us meet these goals, strengthening both public trust and judicial performance.

At IAALS, we believe that the challenges facing the justice system require innovative solutions and continuous improvement. We believe in championing efforts that instill public trust for our society to become stronger. Our work in judicial excellence is setting a new standard of high qualifications, accountability, and impartiality of judges secured through informed election processes. These stronger standards will ensure that communities across the country feel safe in the hands of their civil justice system.

These JPE recommendations represent one example of that new standard. We intend this report to serve as a practical toolkit for states, regardless of their current JPE infrastructure. We hope states will use these recommendations as a roadmap for implementing strategies to modernize judicial practices to support a trusted and trustworthy judiciary.

Brittany Kauffman, CEO

Danielle Kalil, Director of Civil Justice and the Judiciary

INTRODUCTION

Judicial performance evaluation (JPE)¹ plays a crucial role in promoting a fair and effective judicial system by facilitating the assessment of state judges' job performance. JPE programs evaluate judicial performance not based on case outcomes but based on ideologically neutral qualities of impartiality, clarity of communication, legal knowledge, administrative capacity, and judicial temperament.²

How exactly this is done varies by state. JPE programs are often—though not always—integrated within a broader merit selection system for judges, which aims to attract and retain highly qualified individuals for judicial roles. JPE is a key component of a merit selection system, offering a structured evaluation process to support the selection and retention of judges based on their merits, qualifications, and job performance. Merit selection and JPE work hand-in-hand to promote a trusted and accountable system of selection and provide valuable, objective, and relevant information to those tasked with selecting judges. This information is important not only for judges' professional development but also for public accountability and transparency. In some states, JPE serves primarily as a tool for judges to improve their skills. In others, it serves a more outward-facing goal of providing public information about judicial performance, often to voters or others (like state legislators) making decisions about retaining judges.³

While JPE programs vary across states, they generally share the goals of promoting judicial accountability and transparency while maintaining judicial independence.⁴ When it is working well, JPE may contribute to public trust and confidence in courts.⁵ But critics express concern about the accuracy, objectivity, and validity of evaluations and the lack of overall transparency in the process.⁶ Public perception of and relationships with courts are changing, and it is important to modernize the JPE process to ensure it reflects the current needs and expectations of court users, judges, and the public.

The JPE 2.0 project provides a roadmap for modernization. IAALS 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, including through our National JPE Working Group. JPE is a critical component of IAALS' recommendations for promoting effective courts that merit public trust. IAALS has published a number of recommendations for JPE programs over the years, all founded on our research and convening of stakeholders to issue the most up-to-date best practices for evaluating judges.⁷

- ▶ *Transparent Courthouse Revisited: An Updated Blueprint for Judicial Performance Evaluation* (2016)
- ▶ *Recommended Tools for Evaluating Appellate Judges* (2013)
- ▶ *Leveling the Playing Field: Gender, Ethnicity, and Judicial Performance Evaluation* (2012)
- ▶ *Using Judicial Performance Evaluations to Promote Judicial Accountability* (2007)

JPE 2.0, launched in 2021, builds on IAALS' past efforts to develop new best practices for modernizing JPE. This project is not simply about updating recommendations. Rather, it is a response to a rapidly evolving JPE and judicial landscape. The role of judges is changing and is more complex than ever. The number of self-represented litigants has increased exponentially in recent years and continues to grow. Public understanding of JPE and its crucial role in ensuring judicial accountability has waned. Some states have rolled back existing JPE programs. The JPE process itself has become increasingly partisan

and, in some instances, weaponized, undermining its intended purpose. These converging factors underscore the urgent need for a renewed approach to JPE that strengthens public trust in the judiciary, as well as the continued critical need for JPE in our court system.

In enacting this multi-stage JPE 2.0 project, we sought to think creatively about about JPE given this changing landscape, balancing the core goals of judicial performance evaluation with emerging needs, techniques, and concerns about traditional methods. Key steps in the JPE 2.0 project included:

- ▶ **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 JPE Perspectives Survey**, administered to judges in eight states (Alaska, Colorado, Hawai'i, Idaho, Massachusetts, New Mexico, Utah, and Virginia), with findings reported in *National Perspectives on Judicial Performance Evaluation*
 - ▶ **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 stakeholder groups—judges, commissioners, attorneys, and others with a vested interest in JPE—to dive deeper on questions and key issues raised in the JPE Perspectives Survey and the in-person convening
- ▶ **This report of final recommendations and best practices** to improve JPE processes nationwide, informed by the preceding steps

This report presents updated recommendations developed through the JPE 2.0 project, reflecting insights gathered at each stage.

In offering these recommendations, IAALS recognizes that JPE programs vary widely across states, and each state has its own unique culture, considerations, and constraints. Not every state program will feature every element contained in these recommendations. Rather, this report aims to identify and compile best practices that can be adapted and adopted as appropriate for each state. While some of these recommendations are tailored to states with a public-facing component of informing voters about judicial performance, many are broadly applicable to any JPE program. In addition, much of what is contained in these recommendations—for example, content related to the criteria for judicial performance and the role of a judge—have implications for judges and court systems beyond JPE and will be informative and applicable even to states that do not have formal JPE programs.

IAALS remains committed to advancing effective judicial performance evaluation and fostering a more accountable and transparent judiciary across the country. We will continue to offer technical assistance and tools to states seeking to implement these best practices, including guidance for tailoring them to the needs of specific states. IAALS will continue to conduct and disseminate research on emerging issues in judicial performance and explore innovative approaches to evaluation. We will also maintain our role as convener, bringing together stakeholders to share experiences and takeaways, discuss obstacles, and collaborate on solutions. Through these efforts, IAALS aims to ensure every state has access to the tools and knowledge to build a judicial system that earns and maintains the public's trust.

CHALLENGES AND OPPORTUNITIES FACING JPE PROGRAMS

The JPE 2.0 research—including convenings and responses to the JPE Perspectives Survey—revealed a number of recurring challenges related to JPE and many opportunities for improvement. It also underscored the need for a rededication to JPE as critically important to a trusted and trustworthy judiciary. This section summarizes those challenges and opportunities, which provide context for the recommendations that follow.

Declining Public Trust and Confidence in Courts

JPE programs have to contend with the low public trust and confidence that currently exists in institutions, including courts. Partisan factors inevitably affect the judiciary, a fact that JPE cannot ignore. That said, JPE provides an opportunity to promote public trust in the judiciary by seeking public input on judicial performance in a way that is outcome-neutral and not based on political affiliation.

The Changing Role and Expectations of the Judiciary

A recurring theme of the convenings was the changing role of the judiciary and the courts. JPE needs to keep up with these changes. Specifically, there are opportunities to ensure that the criteria used to assess judicial performance reflect the modern role and responsibilities of judges.

Convening participants believed that the current criteria used to evaluate judges are generally the right ones. That said, views on what makes a good judge and a good judiciary have shifted over time as the justice system evolves and the needs of litigants change. As a result, the criteria need to be updated and better defined so they better match the roles, duties, and expectations for judges today. For example, how should judges keep up with new developments in technology? How should they interact with the growing number of self-represented litigants in their courtrooms? In addition to clear definitions, there must be objective and measurable standards, clear rubrics, and uniform application, as the criteria alone can be subjective and difficult to assess.

The Reliability and Limitations of the Evaluation Data

There are concerns that evaluations are too subjective and do not employ an evidence-based process, which may interfere with the reliability of their outcomes. More accurate and more diverse data can provide a fuller picture of judicial performance. In addition, since JPE was developed in the 1980s, we have gained new technological tools and a greater scientific understanding of surveys, data collection, bias, and other factors impacting the reliability of evaluations. There are opportunities to improve how evaluations are conducted by using new tools that did not exist when JPE programs began and applying new information and best practices to elevate JPE.

Surveys are a practical way to gather data about judicial performance, including court users' feedback. At the same time, low response rates are a persistent issue across states and make it difficult to get a representative data set and statistically significant sample size. This means a handful of comments can skew the results, and it contributes to an overall sense among judges that the process is not scientific

or evidence-based. One judge noted in response to the JPE Perspectives Survey that low response rates “do not give a good picture of the performance of each judge.”⁸ Efforts to improve response rates, such as soliciting in-person feedback in the courthouse or providing QR codes for easy access to an electronic survey, have yielded mixed results in practice.

One challenge to robust survey response rates, especially among attorneys, relates to anonymity. Attorneys highlighted the importance of keeping survey responses strictly anonymous amid concern about the potential for judicial retaliation and suggested that they may not answer surveys unless their anonymity is guaranteed. At the same time, withholding all information about respondent characteristics makes it more difficult for judges to put criticism into context and alter their behavior accordingly. One judge described interpreting anonymous comments as “chasing ghosts.”⁹

A related challenge—both for surveys specifically and the process generally—is bias. This is a concern from all perspectives and throughout all aspects of JPE. Unconscious or implicit bias can be reflected in the survey instrument, although tools are available for testing survey questions for bias before they are used. It can also be reflected in the population of survey respondents. Some judges have received inappropriate survey comments (about their race or gender, for example), creating concerns that they are evaluated differently because of their identity and that the data is unreliable. When it comes to mitigating this kind of bias, it can be helpful to have information about who is doing the evaluating (for example, how many times they appeared before the judge). However, it is tricky to collect this information in surveys as it makes them longer and can interfere with anonymity. In addition, there is concern about negative bias—the tendency for people to provide feedback when they are unhappy. There is an opportunity here to develop better practices for addressing bias at all steps of the JPE process, including gathering more objective data and not relying only on surveys.

There were also challenges noted about courtroom observations. Although they are generally considered a helpful tool by judges and evaluators, there is concern that this creates an opportunity for judges to modify their behavior when observed, so observations might not provide an accurate picture of how they run their courtrooms. In addition, capacity to conduct meaningful courtroom observations during the evaluation period is limited, as observations are resource-intensive.

The Inclusion of All Relevant Perspectives in Evaluations

There are opportunities to fill gaps in evaluation data, in part by including important perspectives that have historically been missing from the process.

States are grappling with the challenge of reaching self-represented litigants to incorporate their perspectives. Attorney input is crucial, but many states struggle with getting meaningful attorney involvement in the process (specifically with getting them to complete surveys).

Finally, the perspectives of judges are often overlooked. Judges shared in convenings and in response to the JPE Perspectives Survey 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.”¹⁰ Even if a judge thinks an evaluation result is unfair or information has been taken out of context, they may have limited recourse.

Improper Use of the JPE Process

There are concerns about stakeholders misusing the JPE process. For example, it can be difficult to keep respondents focused on the objective performance of the judge rather than the outcome of the case. Respondents may give poor evaluations because they are unhappy with a ruling.

These concerns extend to the potential for political abuse. For example, members of the public may weaponize the JPE process to remove judges from the bench following a controversial ruling or due to political or ideological disagreements. In addition, some commissioners worry that state legislatures could threaten to withhold JPE funding for partisan reasons or electoral advantage. While concerns that the JPE process may be focused on outcomes rather than the appropriate criteria are not new, the potential weaponization of the JPE process has become an important concern for states and a motivator for reforms.

The Helpfulness of the JPE Process for Judicial Improvement

A central goal of JPE is to help judges improve, but the process is not always conducive to this goal. For example, comments on surveys are not always constructive or actionable. People may be more likely to complete surveys when they are unhappy, particularly about the outcome of a case, which can create a non-representative sample and make responses less helpful. Some comments feature vicious personal attacks, which are unproductive and even harmful. One judge described these as “the equivalent of cyber-bullying” and lamented that there is “virtually no recourse” for this treatment.¹¹ At the same time, there may be truth to even misguided critiques, so it may not always be advisable to dismiss them outright. It can be a challenge to exercise discretion about which comments to pass along to judges. Overall, convening participants shared a belief that comments are more critical today than they have been in the past, and uncertainty remains about how to navigate the tenor of comments.

Judges want feedback, and evaluation is an opportunity for self-improvement. The process of evaluation, including self-evaluation, is valuable. But, for the reasons we have outlined, judges do not always feel the process produces actionable or constructive critiques. One judge shared, “The process does not allow or encourage ongoing feedback or constructive criticism that would allow me to make changes.”¹² And even when feedback is constructive, there are limited resources to translate evaluation results into improved judicial performance, for example through training, education, or mentorship. Many judges find JPE more stressful than helpful, describing the process in the JPE Perspectives Survey using terms like “dreaded,” “degrading,” and “a source of consternation and alarm.”¹³ Judges report that the evaluations negatively impact their mental health and well-being without providing meaningful opportunities to address concerns about their performance.¹⁴ Critical comments can be difficult to receive, and negative evaluations can have serious long-term consequences for a judge’s career, even after they leave the bench. Some judges worry these consequences may deter otherwise interested individuals from becoming judges or seeking retention. Despite the challenges, judges favor receiving feedback, but better supportive structures around judges would help them to receive and integrate this feedback in a more positive and constructive way.

All of this underscores the tension between the dual roles of the JPE process in states where programs aim to both inform the public and support the professional development of judges. On the one hand, transparency to the public about the judiciary’s performance is important. On the other hand, total transparency might hinder a judge’s desire to be completely forthcoming about their own performance and the need to improve. In the JPE Perspectives Survey and convenings, several judges noted that

feedback would be more helpful if it were private and accompanied by mentorship rather than made available to the public.¹⁵ There is room to improve the JPE process to better support judges' performance improvement and to be more intentional about making this process useful for judges while still maintaining the public-facing goals.

Buy-In and Trust of the Process

Although judges generally support JPE, the challenges outlined previously can diminish trust in the process and the results it produces. Judges believe there is little accountability for evaluators and little recourse should they find the process unfair. The JPE process relies on judicial support and buy-in, so it is crucial to build a process that judges trust.

Concerns about buy-in extend to funding and resources. Lack of resources is a barrier to creation and improvement of JPE programs.

Finally, there are concerns about public understanding and trust of the process. While public education and transparency are core components of many JPE programs, resources are limited to accomplish these goals, and often the public does not understand how evaluations work.

Training, Communication, and Education of Stakeholders

The need for better training and education was a recurring theme across convenings and survey responses. The public needs information about the role of a judge, the goals and functions of JPE, and how JPE promotes a trustworthy judiciary. This includes education about the different functions of JPE, judicial discipline, and judicial recusal, and how they interact with one another. Commissioners should receive training on how to conduct fair and constructive evaluations. Judges should receive training about the role of the evaluation, the criteria, how the process works, and how JPE can support their professional development. It can also be helpful for judges to receive guidance on how to keep perspective amid critiques. Attorneys, court users, and other stakeholders should receive information about the importance of their survey completion, their responsibility to take part in the process, and how their feedback is used. All of this is part of a broader opportunity to improve transparency and engagement, which is directly linked to public trust and confidence.

The Role of Technology

A recurring theme was the role technology can play in addressing all these challenges and opportunities, from survey creation and distribution to the mitigation of bias to training. There are opportunities to better incorporate technology at every stage of JPE to modernize the process, get more accurate data, and reach a wider audience.

GUIDING PRINCIPLES

The JPE 2.0 recommendations were guided by ten fundamental principles, which emerged from IAALS' research and work on JPE over the last two decades. These principles seek to modernize JPE programs, account for the needs of stakeholders and the modern role of judges, and improve the process for all involved. In addition to the recommendations that follow, we urge states to consider these principles as well. While the recommendations are designed to be adapted to meet the unique needs of each state JPE program, these guiding principles have applicability beyond formal JPE programs. We encourage states to rely on these principles and the key considerations that follow to guide their work and make the case for robust and effective JPE programs.

- ▶ Recommendations should aim to improve the availability and accuracy of JPE data, including through new and more accurate data collection methods.
- ▶ Recommendations should balance the need for transparency with the recognition that confidentiality in some aspects of the process may encourage judges to more openly receive and act on feedback.
- ▶ Recommendations should promote judges' accountability to court users and the public while also promoting evaluations that do not inappropriately influence judicial decision-making.
- ▶ Recommendations should promote and support judicial performance improvement. Historically, the focus of JPE programs has been the need for reliable data for voters. We now recognize more explicitly that JPE can provide judges with the information and support they need for professional improvement.
- ▶ Recommendations should be ambitious but practical and actionable.
- ▶ Recommendations should promote effective use of time, money, and resources and support programs that can operate effectively within existing constraints and capacity.
- ▶ Recommendations should support the structural integrity and sustainability of JPE programs.
- ▶ Recommendations should provide states with practical tools to modernize JPE and foster a culture of continuous improvement in the judiciary.
- ▶ Recommendations should focus on evaluating what is within an individual judge's control using clearly defined, measurable metrics.
- ▶ Recommendations should account for states' different practices, laws, cultures, and needs.

KEY CONSIDERATIONS

To provide insight into the foundations and logic underpinning these recommendations, we offer the following key considerations. This list reflects the complex interplay of factors that states must carefully weigh and balance when making decisions about JPE. Each state will approach these considerations differently and may need to incorporate additional factors specific to its context. The JPE 2.0 project, in its aim to develop practical and effective recommendations, specifically considered the following needs, realities, and constraints related to the JPE process.

- ▶ The historically solid foundation of JPE as a tool for promoting professional development, informing the public about the work of the judiciary, and reinforcing a high standard of performance by judges.
- ▶ The changing role of the judiciary and the courts within society.
- ▶ The growing numbers of self-represented litigants and the access to justice crisis.
- ▶ The need for trust and confidence in the process from the bench, bar, appointing entities, and the public.
- ▶ The risk of bias in collecting and assessing information about a judge's performance.
- ▶ The vulnerability of the evaluation process to abuse for partisan purposes.
- ▶ The often-complex intersection of JPE, judicial discipline, and judicial recusal, both in the eyes of the law and the eyes of the public.
- ▶ Improved understanding of sources of error in the data collection process.
- ▶ The development of new tools and methods for collecting data relevant to judicial performance.
- ▶ The need for institutional commitment to support strong JPE programs.
- ▶ The benefits of JPE programs collaborating closely with judges while remaining independent of the judiciary.
- ▶ The tension between public transparency and creating an environment conducive to performance improvement for judges.
- ▶ The need to build a culture of feedback and accountability within the judiciary.

RECOMMENDATIONS FOR MODERNIZING JPE PROGRAMS

These recommendations do not prescribe a single “ideal” JPE program. Years of research and practical experience have shown that effective JPE programs are diverse and tailored to the needs of each state. Achieving an effective JPE program involves trade-offs between factors such as cost and transparency, and these trade-offs will vary by state. The ultimate goal of JPE should be a fair, predictable, transparent, and helpful process, and no state will take an identical path to that destination.

By compiling years of research and best practices, the recommendations here present standards and practices that can be adapted to fit each state’s culture and needs.

The recommendations in this report aim to create a future in which:

- ▶ All stakeholders—including judges, court users, and the public—trust that the JPE process is rigorous, credible, and fair.
- ▶ The criteria used to evaluate judges are clear, objective, and reflect the day-to-day jobs of a judge.
- ▶ JPE data is reliable and accurate and reflects a diverse range of sources and perspectives.
- ▶ Judges are equipped with data about their performance as well as resources and support that foster a culture of continuous professional improvement.
- ▶ JPE programs provide data-driven insights to support informed decision-making regarding judicial retention, in states where that is a function of JPE.
- ▶ JPE programs have the resources to support the continual improvement of the judiciary and inform the public.

RECOMMENDATION 1:
Foster trust and confidence in the JPE process.

RECOMMENDATION 2:
Modernize the evaluation criteria to account for the changing roles and responsibilities of the judiciary.

RECOMMENDATION 3:
Implement measures to improve information and reduce errors in evaluations.

RECOMMENDATION 4:
Build a culture of and embrace judicial performance improvement as a key aspect of JPE.

RECOMMENDATION 5:
Display institutional commitment to JPE.

RECOMMENDATION 6:
Ensure and improve transparency for intended audiences.

RECOMMENDATION 1

FOSTER TRUST AND CONFIDENCE IN THE JPE PROCESS.

- 1.1 Design and operate JPE programs in a way that promotes public trust.**
- 1.2 Employ a diverse range of methods to engage the public, increase awareness, and improve buy-in.**
- 1.3 Collaborate with judicial discipline programs to ensure the credibility of information disseminated to the public.**

A key takeaway from the JPE 2.0 convenings and JPE Perspectives Survey is that JPE only works if stakeholders have confidence in the integrity of the process. JPE programs earn confidence through effective design, a clear and reliable evaluation process, and informed stakeholders.

1.1

DESIGN AND OPERATE JPE PROGRAMS IN A WAY THAT PROMOTES PUBLIC TRUST.

Structural Independence of Program

JPE programs must be designed to promote public confidence. One recommended tool for building confidence is for JPE to exist within an independent office that is outside the judiciary and tasked with administration and collection of JPE data. An independent office guards against actual or perceived conflicts of interest and helps the public trust the integrity and impartiality of the evaluation process. While an office outside the judiciary is desirable, it presents challenges to collaboration between the judicial branch and the JPE programs. It is also important that this structural and operational independence does not interfere with accountability of the JPE programs.

If it is not possible to create an independent JPE office outside the judiciary, it should to the degree possible operate independently of the judiciary and courts. This could entail, for example, a separate line-item budget, its own staff, physical separation from court offices, and structural and decisional autonomy.

Structure and Appointment of Commissions

A second recommended tool for building confidence is commissions that are tasked with assessing the JPE data to evaluate judges and provide feedback for improvement. They also recommend retention or removal in states with retention elections. The remainder of this section lays out best practices for commissions, even though commissions are not used in every state and are not strictly necessary for a successful JPE program.

Commission members should be appointed by multiple authorities to enhance real and perceived impartiality, objectivity, and accountability. A majority of commission members should be appointed by entities outside the judicial branch, though judiciaries should be allowed to make a limited number of non-judge appointments. The process of applying to serve on the commission should be open to the public.

To promote stability and institutional knowledge, members' terms should be staggered. Term lengths should be sufficient for members to develop expertise and experience, facilitate the onboarding of new members, and contribute to multiple evaluations cycles. The ideal length will vary by jurisdiction, but generally four to six years is recommended to achieve these goals. To prevent entrenchment, promote accountability, and encourage fresh perspectives, the number of terms a member can serve should be limited. Again, the ideal number of terms will vary by jurisdiction and depend on term length, but a maximum of two to three consecutive terms is generally recommended.

Makeup of Commissions

In the JPE Perspectives Survey, judges shared perceptions that some commissioners have an agenda or bias that interferes with impartial review of their performance.¹⁶ The convenings also highlighted the importance of incorporating a broad array of perspectives in the evaluation process. A balanced and representative commission is crucial for a trusted and trustworthy process for all stakeholders.

To this end, each commission should strive for an equal balance of attorney members and public members who are not attorneys. Within these groups, there should be a diversity of experience and background. For example, attorney commissioners should have diverse backgrounds of experiences to reflect the diverse docket of cases in state courts. It is also essential to ensure that the commission is nonpartisan or evenly divided along partisan lines. This composition mitigates the risk of perceived or actual favoritism or bias based on partisan affiliations. In addition, promoting diversity among commission members, including racial, ethnic, gender, geographic, and ideological diversity, can enhance the depth of perspectives and expand the commission's understanding of the community it serves. Because of their firsthand familiarity with judicial roles and responsibilities, retired judges can make a strong contribution to the commission's work. Because of the potential for conflicts of interest, sitting judges (full-time or retired judges who still sit occasionally) should not serve on evaluation commissions.

The size of the commission should be determined based on the size of the jurisdiction the commission covers, the number of judges to be evaluated, and the practical availability of qualified volunteers.

Training and Rules of Procedure

Training for commissioners and evaluators exists, but this is not always known by judges and attorneys, who may question the expertise and knowledge of the commission. This

was highlighted in the convenings and JPE Perspectives Survey, in which judges shared concerns about lack of training for commissioners as well as potential bias among commissioners.¹⁷ Commission members need support and tools to do their jobs well. All new commissioners should complete training that addresses responsibilities, reviews rules of procedure, familiarizes members with the evaluation process, and gives commissioners tools to recognize and address implicit bias. Additional training should be conducted periodically and as procedural rules are amended. Commissions should have rules of procedure that address the ethical responsibilities of members, including conflicts of interest, political activity, ex parte communications, and confidentiality.

1.2

EMPLOY A DIVERSE RANGE OF METHODS TO ENGAGE THE PUBLIC, INCREASE AWARENESS, AND IMPROVE BUY-IN.

Public Education

Public trust begins with public education. The public generally does not know where to go for reliable information about judges, and there are often misconceptions about what a judge's role is and is not. JPE programs must help the public understand what JPE is, how it works, how they can engage with it, and why it is so important. Public education is not only important for trust and transparency but also for getting feedback about judicial performance.

There are many ways to accomplish this, and JPE programs should tailor public engagement to the needs and culture of their state. For example, a JPE program in a merit selection state may want to help the public understand what merit selection is and the public's role in retaining judges. States use a variety of strategies and media for engaging the public, including public radio, local television ads, and social media. Public education is an ongoing effort that requires the involvement of many stakeholders.

Recommendations to Voters

JPE 2.0 research highlighted concerns that the public may not understand the recommendations provided to voters. Not all JPE programs issue recommendations to voters, but those that do should use clear, objective language. For example, stating that a judge “meets” or “does not meet” minimum performance standards is preferable to “retain” or “do not retain.” This language is more objective and less likely to be perceived as JPE programs directing the public how to vote. It also reinforces to voters that the evaluation process is grounded in predetermined standards applied in that state. Programs can also encourage people to sign up for notifications when JPE recommendations come out.

Media

Education and engagement are an ongoing process that requires the involvement of many stakeholders, including the media.¹⁸ There is an opportunity for media campaigns to help educate citizens about judges, commissions, criteria, and the JPE process. Public engagement and education can encompass a variety of media forums: social media campaigns, radio or podcast advertisements, or commercials on television. Simple, engaging videos can help inform the public.

IN PRACTICE: The Colorado Office of Judicial Performance Evaluation uses short videos posted on their website and social media about what judges do, the public’s role in retaining them, or how to interpret evaluation results.¹⁹

It is important to make information available to the media and others involved in messaging about JPE, including how to interpret and communicate the results. For example, often judges who receive a “does not meet performance standards” result or a “do not retain” recommendation opt not to seek retention and are thus not included in the public report. This may suggest JPE is successfully identifying judges who are not doing their job

well. But this information is not always available or reported. Instead, it can appear to the public that every judge received a “meets performance standards” result or a “retain” recommendation and that the process is not very rigorous. The media needs access to accurate information about JPE results so reporting reflects what is happening and misperceptions do not interfere with public trust. JPE administrators should receive media interview training and build relationships with the media. Funding should be dedicated to publication and promotion of the JPE process, education regarding the process, and availability of JPE recommendations.

1.3

COLLABORATE WITH JUDICIAL DISCIPLINE PROGRAMS TO ENSURE THE CREDIBILITY OF INFORMATION DISSEMINATED TO THE PUBLIC.

The convenings highlighted that while judicial discipline is separate from JPE, this distinction is not always clear to the public. It can be difficult to separate performance from disciplinary matters in public discourse, and often there is overlap between matters of judicial discipline and judicial performance.

While judicial discipline and JPE programs should remain separate entities, they should communicate and coordinate to ensure timely information sharing, particularly when public disciplinary actions are happening simultaneous to a performance evaluation. For example, the judicial discipline process could be expedited for a judge facing retention to allow the JPE evaluation to reflect public disciplinary findings. This coordination is crucial for maintaining public trust in the JPE process. Clarity about the separate roles and functions of these entities, both internally and externally, would also be helpful. Ultimately, the goal is for the two processes to work together to promote public trust in the judiciary.

RECOMMENDATION 2

MODERNIZE THE EVALUATION CRITERIA TO ACCOUNT FOR THE CHANGING ROLES AND RESPONSIBILITIES OF THE JUDICIARY.

- 2.1 Use fair and measurable criteria reflective of the modern role and responsibilities of a judge.**
- 2.2 Evaluate judicial competence broadly, to include more than legal knowledge.**
- 2.3 Evaluate a judge's ability to be impartial, both with respect to personal beliefs and with respect to the treatment of litigants, lawyers, and other court participants.**
- 2.4 Evaluate a judge's clarity of communication, including all communication from the judge's chambers.**
- 2.5 Evaluate judicial temperament in a way that focuses on specific and observable behavior.**
- 2.6 Evaluate administrative capacity in a way that is aligned with the judge's actual role and responsibilities and assesses factors within the judge's control.**
- 2.7 Evaluate a judge's efforts to advance and improve the justice system.**

While specific criteria vary by state, historically, JPE has evaluated judges on the core criteria of legal knowledge, impartiality, clarity of written and oral communication, judicial temperament, and administrative capacity. These criteria are meant to evaluate the process of judging, not the outcomes of individual cases.

The JPE 2.0 convenings underscored the need to reassess the current evaluation criteria in light of the evolving role of judges and the changing landscape of the justice system. Participants noted that conceptions of justice and expectations for judges have evolved significantly since the criteria were developed in the 1980s. For example, judges require greater technological proficiency, and the rising number of self-represented litigants has shifted how judges must engage with parties and run their chambers and courtrooms. Furthermore, the original criteria were developed by the ABA and reflect attorneys' views about what makes a good judge. The criteria need to be updated to reflect other views, such as those of litigants.

While participants generally agreed that the traditional criteria remain relevant, they also identified areas for modernization and the need for clear definitions and measurable standards. This section presents recommendations to enhance assessment criteria, consistent with the insights from the convenings.

Evaluating case outcomes is expressly not a part of JPE. Rather, JPE is designed to evaluate the process of judging. Evaluating judges based on decisions in specific cases could undermine judicial independence. Judges must be free to reach decisions based on their interpretation of the law, even if that results in an unpopular or controversial opinion. Evaluation of case outcomes could incentivize judges to rule in ways that are popular or politically expedient to avoid poor evaluations, which could compromise their impartiality. JPE assesses objective aspects of judicial performance and is not an appropriate avenue for disagreements about substantive legal decisions. That is the job of the appeals process.

2.1

USE FAIR AND MEASURABLE CRITERIA REFLECTIVE OF THE MODERN ROLE AND RESPONSIBILITIES OF A JUDGE.

The goal of JPE is to assess a judge's performance with respect to the process of judging using neutral measures and predetermined benchmarks. The criteria for judges' performance must be objective, measurable, separated from case outcomes, and clearly articulated.²⁰

To achieve this, criteria should measure aspects of judicial performance within a judge's direct control. For example, a judge does not control the layout of the courthouse or the behavior of court staff and should not be assessed on those elements of a court user's experience. The judge does control how they treat people who come before them and how clearly they communicate procedures, so they should be assessed on those elements.

In addition, performance standards should reflect the day-to-day realities of various judicial roles, for example, accounting for differences in trial versus appellate roles and problem-solving court roles, as well as the types of cases and the types of litigants the judge typically encounters.

Finally, JPE programs should clearly define each evaluation criterion and ask specific questions about observable behaviors. This will help ensure consistent interpretation and avoid ambiguity.

IN PRACTICE Instead of broad questions about whether a judge is "professional," evaluation tools should provide specific questions and examples of professional behavior. Did the judge appear bored with the case or inattentive to what was happening? Was the judge unprepared? Did the judge start court on time? Did the judge address individuals with a disrespectful tone?²¹

Based on the research and perspectives outlined previously regarding the need to modernize assessment criteria, we recommend evaluating judicial performance using the following criteria:²²

- ▶ Judicial competence
- ▶ Impartiality
- ▶ Clarity of communication
- ▶ Judicial temperament
- ▶ Administrative capacity
- ▶ Efforts to advance the justice system

These criteria build upon existing standards, incorporating feedback and research from JPE 2.0 to introduce important updates and a new category. The next sections offer updated guidance for each criterion.

2.2

EVALUATE JUDICIAL COMPETENCE BROADLY, TO INCLUDE MORE THAN LEGAL KNOWLEDGE.

JPE has historically evaluated a judge's legal knowledge or legal ability, defined as legal reasoning ability; knowledge of substantive law; knowledge of rules of procedure and evidence; and keeping current on developments in law, procedure, and evidence.²³ While legal knowledge is crucial, it represents only one facet of a judge's overall competence and abilities. We recommend replacing the traditional criterion of "legal knowledge" with "judicial competence" to encompass a broader range of skills, including the following:

- ▶ **Legal reasoning.** Does the judge demonstrate the ability to apply laws, rules, and precedent to facts to reach logical and legally sound conclusions?
- ▶ **Knowledge of substantive law.** Does the judge display an understanding of substantive law, including keeping up to date on current developments in the law?

- ▶ **Knowledge of rules of procedure and evidence.** Does the judge display an understanding of the rules of procedure and evidence, including keeping up to date on current developments?
- ▶ **Competence with technology.** Does the judge know how to use technology as needed to discharge their judicial duties, including keeping up to date on current developments?
- ▶ **Knowledge of the community.** Does the judge have an understanding of the community and the people who will appear before them? Does the judge have knowledge of local resources?
- ▶ **Context-specific application.** Does the judge understand and apply the law, tools, and practices appropriate to the specific needs of their court, docket, and parties?

Judicial competence may look different for different judges. For example, there is a steep learning curve for new judges, and evaluations should account for the fact that a judge’s knowledge varies over the course of a judicial career.

2.3

EVALUATE A JUDGE’S ABILITY TO BE IMPARTIAL, BOTH WITH RESPECT TO PERSONAL BELIEFS AND WITH RESPECT TO THE TREATMENT OF LITIGANTS, LAWYERS, AND OTHER COURT PARTICIPANTS.

The concept of judicial impartiality is not new, but it remains a cornerstone of the justice system and of public trust in courts. JPE programs should continue to assess judges’ impartiality, considering factors such as treating everyone with dignity and respect, avoiding improper bias, considering all sides of an argument, and making difficult decisions. Additionally, JPE programs should evaluate whether judges exhibit bias or prejudice based on race, sex, religion, national origin, disability, age, sexual orientation, political affiliation, socioeconomic

status, legal representation, or mental health. They should also assess whether judges apply the law fairly and impartially.

2.4

2.4 EVALUATE A JUDGE’S CLARITY OF COMMUNICATION, INCLUDING ALL COMMUNICATION FROM THE JUDGE’S CHAMBERS.

This criterion has traditionally been termed “clarity of written and oral communication.” “Clarity of communication” is a more fitting descriptor and pertains not only to decisions but to factual findings, orders, opinions, etc. It should encompass the following categories of clear and logical communication: verbal communication in court; written decisions; and written and oral communications about court procedures, processes, and policies throughout the life of a case. It should also encompass adaptation of communication styles for specific audiences, including the use of plain language.

Judges should ultimately be responsible for the clarity and accuracy of communications issued from their chambers, whether written by the judge, a law clerk, an intern, a staff member, or even artificial intelligence, and evaluations should reflect the exercise of that responsibility. That said, it is important to separate communications by the judge’s chambers from communications by other court entities, like the clerk’s office, for which the judge is not directly responsible.

2.5

EVALUATE JUDICIAL TEMPERAMENT IN A WAY THAT FOCUSES ON SPECIFIC AND OBSERVABLE BEHAVIOR.

Judicial temperament speaks to the way the judge manages proceedings, behaves in and out of the courtroom, and treats people with whom they interact in their role. The definition of judicial temperament should encompass the following:

- ▶ Acting with patience and self-control, remaining composed under pressure
- ▶ Treating people with appropriate courtesy, respect, dignity, and compassion in all encounters, including through written and electronic communications
- ▶ Promoting understanding and clarity for all parties involved, including self-represented litigants
- ▶ Maintaining control over court proceedings
- ▶ Displaying attentiveness to the proceedings and to the parties
- ▶ Exhibiting sensitivity to cultural differences and interacting effectively with people from different cultures

This criterion encompasses a judge’s behavior not only in the courtroom but also in remote hearings, in chambers, and in written communications. What constitutes appropriate judicial temperament may differ depending on the context.

Judicial temperament can mean different things to different people, so JPE programs need to clearly define what they mean. The description of this criterion should include examples of specific and observable behavior to make clear what it entails, and questions should ask about specific behaviors rather than general opinions about the judge’s temperament.

IN PRACTICE Instead of broad questions about a judge’s temperament, evaluation tools should provide specific questions and examples of appropriate temperament. Did the judge overreact to incidents in the courtroom? Did the judge appear bored with the case? Did the judge appear attentive to what was happening in the courtroom? Did the judge make eye contact with the people they spoke to? Did the judge use intimidation to maintain control of the courtroom?²⁴

2.6

EVALUATE ADMINISTRATIVE CAPACITY IN A WAY THAT IS ALIGNED WITH THE JUDGE’S ACTUAL ROLE AND RESPONSIBILITIES AND ASSESSES FACTORS WITHIN THE JUDGE’S CONTROL.

Administrative capacity captures a judge’s efforts to promptly and fairly resolve cases, contributing to a more just and accessible experience for court users. Assessment of a judge’s administrative capacity should be aligned with the unique role and responsibilities of each judge, including their position (i.e., whether they serve in a chief or presiding role), docket needs and case types, and litigants in their courtroom. Administrative capacity should encompass the following:

- ▶ Deliberate time management, both in and out of the courtroom, that facilitates effective processes and scheduling as well as timely rulings that keep a case moving toward resolution
- ▶ Proactive case management, including early intervention; clear expectations and timelines; regular case management conferences and monitoring; streamlined procedures; and encouraging cooperation among parties
- ▶ Effective use of technology to manage cases, filings, and proceedings
- ▶ Effective use of innovative tools or processes as necessary

This criterion should not encompass any aspects of administration outside of the individual judge’s direct control (i.e., clerk’s office procedures, mandates from the state administrative office or state supreme court, rules of evidence).

2.7

EVALUATE A JUDGE'S EFFORTS TO ADVANCE AND IMPROVE THE JUSTICE SYSTEM.

We recommend the addition of a new criterion that evaluates a judge's efforts to advance and improve the justice system. Some states already evaluate a judge's community involvement. The addition of this criterion builds on the engagement targeted by community involvement but encompasses a broader set of activities.

Community involvement is not included in the current ABA criteria. However, some states include it because they believe a judge's involvement in the legal community or the local community is relevant to their judicial performance. In addition, it gives evaluators a more holistic view of a judge. States have different ways of assessing and weighing this criterion. For example, in Colorado, judges are evaluated on their "service-oriented efforts designed to educate the public about the legal system and improve the legal system."²⁵

There are some critiques of including community involvement in a judicial evaluation. The first is that community involvement is not a core aspect of judging, and it is unfair to expect judges to go above and beyond their day-to-day job. A judge can be great at their job without going out into the community. Thus, some argue this is not an essential criterion. The second critique is that capacity for community involvement differs among judges. For example, new judges may need to devote more time to learning the job, whereas more experienced judges may have more time to contribute to other efforts. This makes it difficult to evaluate community involvement fairly. The third critique is that judges may have good reason to maintain separation from the community, for example due to ethical or security concerns. Finally, this criterion may interfere with anonymity. In some states, like Missouri, evaluators review judges without knowing each judge's identity in an effort to minimize the risk of bias. In states that take this approach, the inclusion of contributions

to the community may make it too easy to uncover the identity of a judge.

To address these critiques, we recommend a new criterion that evaluates a judge's efforts to improve the justice system and/or the administration of justice. The modern role of a judge extends beyond mere adjudication of cases and includes taking steps to promote the efficacy and fairness of the courts. This new criterion aims to reflect this reality without placing an unduly high burden on judges. It addresses the critiques outlined previously by encompassing a broader definition of "community" that includes the broader legal community, not just the judge's local community. It also encompasses a broader set of activities, including efforts to improve the system or existing practices, committee service, community and public education, legal system innovation, and mentorship. This can entail work at the local level, the state level, or the national level. By participating in these efforts, judges contribute to a more accessible and equitable legal system and improve trust and confidence in the courts.

To effectively evaluate this criterion, states should establish clear, objective, and ideologically neutral metrics, providing specific examples of activities that demonstrate a judge's commitment to improving the justice system. That said, the specific activities will vary greatly depending on the judge, their experience level, and the resources available, so evaluation of this criterion should be flexible and consider those unique circumstances. There need not be a minimum time commitment required, but judges should be able to demonstrate a proactive effort beyond the adjudication of individual cases.

Typically, criteria are weighed equally. It may be appropriate to weigh this criterion differently from the others. The purpose of this criterion is not to make the job of a judge more difficult or burdensome but rather to convey that the role of a judge extends beyond their caseload and into the legal system and broader community.

RECOMMENDATION 3

IMPLEMENT MEASURES TO IMPROVE INFORMATION AND REDUCE ERRORS IN EVALUATIONS.

- 3.1 Use evidence-based practices to ensure surveys are fair and generate reliable data.**
- 3.2 Gather a broad range of data using new and diverse evaluation tools.**
- 3.3 Take steps to increase response rates to surveys to improve reliability of data.**
- 3.4 Assess performance regularly, and allow judges to respond to inaccuracies.**

The JPE Perspectives Survey and convenings highlighted opportunities to modernize approaches used for evaluations. Evaluations must be based on accurate information to be fair and useful to judges and deserving of the public's trust, and JPE must use reliable technology and methodology to collect that information. There have been significant technological and scientific advances in survey collection and methodology since JPE was first developed, as well as new insights on gathering user perspectives on our justice system. The recommendations in this section outline ways to improve the accuracy and reliability of data and tools used in evaluations, ensuring they reflect modern methods.

3.1

USE EVIDENCE-BASED PRACTICES TO ENSURE SURVEYS ARE FAIR AND GENERATE RELIABLE DATA.

Most JPE programs rely on surveys to some extent. They are a necessary tool to gather feedback about a judge's performance from key court users. Despite widespread agreement on the importance of surveys, the JPE 2.0 research highlights the critical need to enhance the reliability of survey data. Specific concerns include low response rates, the subjectivity of feedback, and the potential for bias. The recommendations that follow focus on improving the reliability of survey feedback.

Surveys in Context

While surveys are a valuable component of the evaluation process, they should not be the only evaluation tool. Rather, they should be one source of data about a judge's performance, to be paired with independently collected objective data. (Additional data points are discussed later in this section.) Diversifying evaluation methods can make assessments more objective and lessen the impact of some of the shortcomings of surveys.

While each JPE program will choose how much weight to assign to survey responses, states should be mindful not to overweigh surveys. If survey data is internally inconsistent, or inconsistent with other information about the judge (i.e., courtroom observation or docket data), further investigation is warranted.

Survey Respondents

JPE programs should survey many types of respondents, not just lawyers. Examples of other categories of survey respondents include litigants (including those who are self-represented), court staff, social workers, law enforcement, probation officers, jurors, court reports, peer judges, and appellate judges. Research suggests that a

wide-ranging respondent base is important not only for gathering reliable feedback, but also for mitigating bias in survey responses.²⁶

Regardless of the nature of their interaction, only those who have *recently* interacted with judges and can speak to their recent performance should be invited to complete surveys. It may be appropriate to give different weight to survey responses based on the nature and extent of the respondent's interactions with the judge. Screening questions can help ensure surveys are only being completed by individuals with personal knowledge of the judge. That said, anonymity of respondents is critical to gathering honest feedback, and screening questions have the potential to interfere with anonymity. Program administrators should balance the usefulness of survey responses with the need for anonymity.

IN PRACTICE Rather than asking the exact number of times a respondent appeared before a judge, they could ask if a respondent appeared before that judge just once or multiple times. They should also remove any potentially identifying information from survey responses before sharing them with an evaluated judge.

Furthermore, JPE programs should educate survey respondents, especially attorneys, about the confidentiality of surveys to mitigate concerns about retaliation and facilitate their participation.

IN PRACTICE The first page of the survey could include language or a link to a video explaining that the surveys are confidential, how data is stored, and what the judge will get to see. Some states offer attorneys CLEs on judicial performance evaluation, and explaining confidentiality is one component of those sessions.

Survey Questions

Both the convenings and JPE Perspectives Survey emphasized the need to design surveys that promote constructive feedback and discourage personal attacks.²⁷ To achieve this, survey questions should focus on specific, readily observable behaviors rather than general attributes and ask respondents to recall specific examples of those behaviors. Using a frequency-based response scale ("never" to "every time") or a multiple-point Likert scale (from "strongly agree" to "strongly disagree") is preferable to subjective grading or binary choices (i.e., "acceptable" or "unacceptable"). This can improve data reliability and mitigate bias.

To further mitigate bias, surveys can incorporate reverse coding and reflective questions to promote thoughtful responses. Additionally, structured free recall can be used. This technique prompts respondents to recall specific instances of behavior before completing the survey. This has been shown to improve the accuracy and completeness of information retrieved and minimizes the potential for respondent bias.²⁸

IN PRACTICE A structured free recall section at the beginning of a survey could ask respondents to recall some positive and negative behaviors they have observed in the judge they are evaluating.

Survey questions should be tailored to different audiences based on their respective experiences with the judge and their ability to assess particular criteria. This makes it more likely that the feedback collected is relevant and actionable.

IN PRACTICE Attorney surveys may feature questions about legal knowledge, while court staff surveys feature questions about administrative capacity, and litigant surveys feature questions about the ease of navigating the court process before that judge.

In the JPE Perspectives Survey, judges expressed difficulty acting on feedback when they lacked information about who is providing it, such as their role or level of interaction with the judge.²⁹ At the same time, attorneys insist on anonymous surveys. As long as the confidentiality of respondents can be maintained, attorney surveys should begin with demographic questions about the nature of the attorney's practice and interactions with the judge (i.e., attorney's years of practice, percentage of practice devoted to civil or criminal cases, frequency with which the lawyer appears in court). This gives judges and the commission a sense of the extent to which respondents are representative of all potential respondents, regardless of response rate. Responses should be reported in the aggregate, not associated with specific questions, to protect anonymity.

Mitigating Bias

Bias in survey responses emerged as a persistent concern across all convenings and the JPE Perspectives Survey.³⁰ There is a widespread perception that women and minority judges may be evaluated more harshly due to implicit or explicit bias of survey respondents. There is also a fear that political affiliation or ideology influences survey responses. There is also concern about "negative bias," the possibility that survey respondents may be more likely to complete surveys when they are dissatisfied with a judge, potentially skewing the results.

Beyond the design of survey questions, there are other steps JPE programs can take to reduce the opportunity for bias. For example, some states have begun offering CLE credit to attorneys for implicit bias training, which can be completed immediately before the evaluations. To ensure surveys are using best practices, they should be developed in consultation with experts in job performance evaluations and survey design. It is impossible to eliminate bias entirely, even when surveys are designed with that in mind, so additional safeguards are necessary once survey

feedback is compiled. For example, JPE programs should examine if respondents systematically favor judges of a certain race, gender, or ethnicity and, if so, should take steps to redesign the survey.

In addition, it is crucial for evaluators to contextualize survey responses. For example, if a judge receives five critical responses out of ten surveys but presided over 1,000 cases that year, this provides valuable context for interpreting these survey results. This approach helps to mitigate the "negative bias," which disproportionately focuses on isolated criticisms, while still acknowledging and addressing concerns raised by court users.

Screening Inappropriate Comments

While judges generally find written comments helpful, JPE 2.0 research highlighted the need for more careful screening of comments. It also highlighted the challenge for commissioners of determining when and how to define and filter out inappropriate content. Judges have reported receiving distressing comments, including threats and personal attacks, which offer little value when it comes to performance improvement.

JPE programs should develop a clear process for removing threats, personal attacks, and comments unrelated to judicial performance from completed surveys. This includes comments motivated by personal animosity due to adverse rulings and inappropriate comments based on the judge's demographic characteristics. Reviewers must exercise discretion in identifying and removing inappropriate comments and balance this with the need for transparency and nuance. Surveys should explicitly state that threats and inappropriate comments will be removed and clearly describe what constitutes an inappropriate comment. States have found that this messaging reduces the number of inappropriate comments. In addition, providing instructions about how to provide constructive feedback and offering implicit bias training at the beginning of surveys may enhance the quality of comments.

Timing and Frequency of Surveys

The timing and frequency of survey distribution and completion is important but tricky to get right. If surveys are sent infrequently, there may not be enough data. If surveys are sent too frequently, respondents may experience survey fatigue. If surveys are sent too soon after a hearing, respondents may be motivated by strong emotions about a case outcome rather than a thoughtful assessment of the judge's performance. If surveys are sent too long after a hearing, respondents may have difficulty recalling specific details or may lose interest.

Rather than administering the survey at regular intervals, JPE programs should consider the needs of different groups and when people are surveyed relative to the timing of their case. For attorneys, timing is not as critical, as they are more likely to have multiple appearances before a judge. For litigants, survey responses can be improved if they are completed closer in time to their interaction with the judge. It is recommended that self-represented litigants be surveyed at the conclusion of their case. Surveying them earlier might cause worry that their survey responses, or even their willingness to participate, could potentially influence the outcome of their case.

3.2

GATHER A BROAD RANGE OF DATA USING NEW AND DIVERSE EVALUATION TOOLS.

The JPE 2.0 research revealed the desire for more data generally, as well as more objective data. A JPE program should gather a broad and deep set of information on the judge's performance that is timely and based on objective data or the personal knowledge of court users. This should include a variety of data sources: survey data from a variety of stakeholders, review of case management practices and written opinions, courtroom observation, and interviews with the judge. The breadth of data provides a fuller, more accurate picture of a judge's performance and can mitigate the impact of potential personal or negative biases

in surveys by offering additional (and sometimes more objective) data points. Next are recommendations about specific evaluation tools.

Case Management Data

Case management data can provide an objective snapshot of aspects of a judge's performance. When assessing case management data, it is important to focus only on aspects over which the judge has direct control. The following metrics are examples of case management data that could be used: time to disposition, clearance rate, total docket numbers, and length of time to rule on motions.³¹ Evaluators can also ask judges about their case management approach during an interview.

Courtroom Observations

Courtroom observations allow evaluators to see how a judge runs their courtroom and interacts with court users. Judges surveyed in the JPE Perspectives Survey generally valued the input of courtroom observers, viewing them as unbiased parties not personally invested in case outcomes. However, some noted a need for improved observer training.³² Courtroom observers should receive training, templates, and rubrics to guide their observations and to ensure judges are assessed on the same qualities and skills as their peers.

The convenings noted a concern that judges might alter their behavior when being observed, or that the short duration of observations might not provide a complete picture of typical courtroom practices. To mitigate these concerns, it is ideal for judges to be observed on separate occasions by at least two different observers during the evaluation period, but this is not always possible. To increase capacity, JPE programs can consider piloting volunteer programs with law students, paralegals, graduate students, or other trained volunteers as observers and tapping into local networks to observe locally.³³ To minimize the opportunity for judges to alter their behavior during observations, JPE programs can consider live virtual observation or reviewing recorded courtroom sessions.

Decision and Opinion Review

Evaluation of judges should include review of the legal reasoning and analysis, fairness, and clarity of their decisions and opinions, without regard for outcomes. Each judge should select a few representative decisions and opinions, ranging in case type and complexity. This can include written decisions and opinions or oral ones that have been transcribed.

Ideally, the review should be conducted by teams composed of at least one attorney and one public member who is not an attorney. This approach ensures a manageable workload and allows for a more thorough and nuanced review. It also enables evaluators to assess whether the judge's decision or opinion is easily understood by someone without legal training. If the commission includes former judges, their expertise can be valuable for consulting on matters that would benefit from a judicial perspective.

To ensure consistent application of the criteria, evaluator training is important. Training should explain evaluation criteria and emphasize the importance of focusing on objective process rather than case outcomes. It should also educate public members about the functions of appellate courts and the opinion-writing process. A rubric or template can facilitate consistency and fairness of evaluations.

It has been suggested that JPE programs should consider a judge's reversal rate as an objective measure of performance. While this may seem straightforward, the reality is much more complex. Merely calculating a reversal rate presents a confusing and incomplete picture. Rulings are reversed for many reasons, not all of which indicate an issue with a judge's performance. Evaluators would need to determine not just the rate of reversal but the source of each appeal and the reason for reversal, which is a complicated task that requires a significant investment of resources. Also, the individual judge's reversal rate is meaningless without the proper context;

for example, comparing a judge's reversal rates to historic rates or those of other judges.³⁴ Making these assessments is incredibly difficult, resource-intensive, and often impractical. It is simply too difficult to develop a methodology that is fair, consistent, and results in something the public can easily understand. Ultimately, while reversal rates may seem like a simple data point, they are not very helpful metrics of good judicial performance. Instead of focusing on reversal rates, JPE should prioritize other, more valuable data points. For example, evaluators can request that judges provide an opinion that has been reversed for review. Evaluators can then request the appellate decision to gain insight into the reasons for the appeal and assess whether it indicates a performance issue.

Self-Evaluation

Self-evaluation allows judges to offer valuable context and present a more complete picture of themselves to evaluators. It also offers judges an opportunity to engage with the process, reflect on their performance, and take ownership of their professional development. Self-evaluations need not be shared as part of the evaluation report, as privacy may encourage more honest self-reflection. Self-evaluations can take many forms (i.e., survey, interview), and JPE programs should build their self-evaluation in whatever way works for their state and facilitates reflection most effectively.

Public Input

JPE programs should give the public an opportunity to weigh in about the performance of individual judges, either through public hearing or anonymous written comment. These opportunities promote transparency about the JPE process and convey that the judiciary cares about performance and public opinion. Opportunities for input should be publicized widely and well in advance.

As with surveys, opportunities for public comment should include an explicit statement that threats or inappropriate comments will not be tolerated.

Members of the public making threatening or inappropriate comments in a public hearing should be removed, and JPE administrators should comb through anonymous written comments and take out any threatening or inappropriate comments before sharing them with judges. To help the public participate meaningfully, provide them with information about the evaluation criteria when soliciting feedback. This can help members of the public understand how judges are evaluated and tailor their comments to make them more useful to both evaluators and judges.

Interviews with Judges

Judges should have an opportunity to participate in the evaluation process. The completed evaluation report should be shared with the evaluated judge and followed by an interview in which the judge can respond to the content of the report and the evaluators can help interpret evaluation results and discuss areas of concern. Evaluators should provide judges with enough information about their evaluation to facilitate their engagement with the interview, as it is challenging for judges to respond when they do not have complete information.

3.3

TAKE STEPS TO INCREASE RESPONSE RATES TO SURVEYS TO IMPROVE RELIABILITY OF DATA.

Survey Distribution

Low response rates were a nearly universal issue identified in both the stakeholder convenings and JPE Perspectives Survey.³⁵ This challenge, common across states, creates a perception that data is unreliable, makes it hard to assess whether there are patterns (and if so, what they are), and can give individual comments disproportionate weight in influencing outcomes.

States should prioritize strategies to improve survey response rates, including improved survey distribution methods. QR codes, electronic

distribution, text message notifications, and similar techniques can help. Surveys conducted electronically, with appropriate security controls, can both cut costs and improve response rates and accessibility. To facilitate electronic surveying, JPE programs should work with court administration to acquire email addresses or other electronic contact information for as many users as possible and include a clear message that surveys are safe to open. Including an official court or government seal, as well as contact information for someone who can answer questions, can enhance credibility and trustworthiness. Programs should explore ways to address cybersecurity concerns such as information on what is safe to click and use of an email address from a recognizable and trusted source.

Efficient survey distribution methods should be combined with improved and shortened surveys to increase response rates. Someone with authority to encourage responses should remind people in the courtrooms that they will likely receive surveys, that they are encouraged to fill out the survey, and that their feedback is welcomed.

IN PRACTICE Utah has sent reminders using a pre-recorded message by the state's chief justice. In other states, JPE administrators encourage judges or clerks to remind litigants to complete surveys.

Messaging

To increase buy-in to the JPE process and boost participation, JPE programs should tailor survey messaging for each audience, explaining the value of their feedback and how it contributes to a better judicial system. For example, litigants, including self-represented litigants, can offer helpful feedback about the accessibility of court processes. Attorneys can offer valuable insight into a judge's legal knowledge and analysis. In addition, messaging to attorneys can call on their ethical duties, framing surveys as part of attorneys' duties

as officers of the court and reminding them against making a false statement about a judge. Judges or court staff should consider reminding those with whom they interact that they will likely be receiving surveys, they are encouraged to fill them out, and their feedback is welcomed.

Self-Represented Litigants

JPE programs should make an effort to capture the opinions and experiences of self-represented litigants (SRLs). The percentage of cases with self-represented litigants on at least one side has increased exponentially in recent years, and this trajectory is likely to continue. It is well-documented that at least one party is self-represented in over 70% of civil cases. More than ever, the job of a judge involves effectively interacting with self-represented litigants, which is a shift from the traditional role of interacting with attorneys on both sides. Yet the perspective of this group remains underrepresented in performance evaluations. In fact, their feedback is largely uncollected. This is a significant area for improvement and innovation in JPE, as self-represented litigant voices need to be brought into the process. These litigants have valuable things to say about the judges they interact with, especially whether they understood what happened in their case, a crucial aspect of a judge's effectiveness in interacting with self-represented litigants.

Accomplishing this is a challenge for JPE program across the country, with states trying different tactics and studies being undertaken to better understand how to capture the experiences of self-represented litigants. Efforts like court kiosks and QR codes have yielded mixed results. Reaching self-represented litigants will require a multi-modal approach to outreach and distribution, encompassing both in-person and electronic efforts. Given the developing nature of this endeavor, JPE programs should pay attention to emerging tactics and strategies and be ready to put them in place.³⁶ JPE programs should work with

court administrators to gather contact information for self-represented litigants, specifically email addresses and cell phone numbers. General recommendations for improving survey response rates apply for this population, as well: explaining why the SRL perspective is crucial, making surveys brief and easily accessible, and distributing them widely.

3.4

ASSESS PERFORMANCE REGULARLY, AND ALLOW JUDGES TO RESPOND TO INACCURACIES.

Sitting judges should be evaluated on a regular schedule. Regular evaluations promote predictability and accountability for judges by establishing clear expectations. The frequency of evaluations will depend on the needs and features of each state. For example, some programs conduct an interim or initial evaluation that is internal only. Judges should have a chance to review their assessments and challenge aspects they deem inaccurate or unfair before they are made public. If the judge's concerns are not considered in revising the report, the judge should be permitted to make a statement in or adjacent to the publicized report.

RECOMMENDATION 4

BUILD A CULTURE OF AND EMBRACE JUDICIAL PERFORMANCE IMPROVEMENT AS A KEY ASPECT OF JPE.

- 4.1 Incorporate professional development into messaging about the objectives of JPE.**
- 4.2 Engage judges and the judicial branch throughout the evaluation process, promoting the use of resources, training, and support to enhance judicial performance.**
- 4.3 Educate judges about the value of JPE for their own professional development.**

Across states, JPE plays the dual roles of informing the public and enhancing judicial performance. In some states, JPE serves the sole purpose of professional development for judges. In other states, the primary focus on informing voters can overshadow the goal of improving judicial performance. Though these goals can at times be in tension with one another, both are crucial components of an effective JPE system, and performance improvement should not be ignored. The JPE Perspectives Survey and convenings revealed that judges value feedback and consider the JPE process beneficial to their professional growth.³⁷ For many judges, JPE evaluations are the only time they receive feedback about their performance. At the same time, they want more actionable feedback and more robust resources to support its implementation.

4.1

INCORPORATE PROFESSIONAL DEVELOPMENT INTO MESSAGING ABOUT THE OBJECTIVES OF JPE.

Even states that use JPE primarily for the purpose of voter education should make clear that evaluations are not just a tool for retention but also a critical tool to help judges improve. This messaging is important for all JPE stakeholders, including judges, evaluators, and the public. Performance evaluations are commonplace in most jobs, and judges should be no different. Even judges who are performing well can benefit from constructive feedback.

4.2

ENGAGE JUDGES AND THE JUDICIAL BRANCH THROUGHOUT THE EVALUATION PROCESS, PROMOTING THE USE OF RESOURCES, TRAINING, AND SUPPORT TO ENHANCE JUDICIAL PERFORMANCE.

Self-improvement is an important part of JPE, but a key challenge identified in the convenings and JPE Perspectives Survey was the lack of support for such improvement.³⁸ Building a culture of performance improvement requires more than just a shift in messaging. Judges must have the support, resources, and time they need to improve.

JPE programs should collaborate with judiciaries and judicial education offices to promote opportunities for judges to improve their performance following evaluations, regardless of the judge's experience level. JPE should be one component of a culture in which professional growth is normalized for judges. This culture should be built around the knowledge, skills, and abilities needed for quality judges. JPE programs have information that judicial education coordinators or mentor judges should know, and they should collaborate as much as possible with the judiciary to support professional development by providing this information.

Judges should have the benefit of personalized professional development plans to help them incorporate the feedback received, and some states have found the use of mentor or facilitator judges helpful. JPE should work with offices of judicial education to facilitate the creation of these plans. Ideally, the process would include goal setting by the judges, and programs should look to recommendations from experts on professional development. JPE can offer useful information to help the judiciary provide support for judges whose performance needs targeted improvement, such as mentorship or coaching. Support should be individualized to each judge's performance evaluation and should be offered soon after results are shared. It can be useful to offer new judges an interim evaluation before they are first subject to retention to promote opportunities for self-improvement.

Judicial wellness support throughout the JPE process is also important,³⁹ including context for negative or personal comments, as judges report it can be distressing to review negative comments, especially without any support or context.

4.3

EDUCATE JUDGES ABOUT THE VALUE OF JPE FOR THEIR OWN PROFESSIONAL DEVELOPMENT.

While continual improvement is a normal aspect of professional life, judges have few opportunities for honest feedback about their work. JPE provides such an opportunity, and judges should be educated about the potential benefits to them of JPE. By reflecting on their strengths and areas for growth, judges can adjust their practices to better meet the needs of court users and improve their interactions with litigants. This not only fosters public trust in the judiciary but also supports the ongoing professional development of individual judges.

Education alone is not enough to advance a culture of embracing professional development. Judges must trust the evaluation process and have access to resources and support to implement feedback effectively.

RECOMMENDATION 5 DISPLAY INSTITUTIONAL COMMITMENT TO JPE.

- 5.1 Establish JPE programs that are enduring.**
- 5.2 Provide programs with the resources needed to effectuate the goals of JPE.**
- 5.3 Seek buy-in and cooperation from all stakeholders.**

JPE 2.0 convenings emphasized the need for strong institutional support to enable JPE programs to achieve their objectives. States must demonstrate commitment to JPE through adequate funding, adequate staffing, and building a culture of accountability and improvement. Strong leadership and champions are essential for the success of JPE programs.

5.1

ESTABLISH JPE PROGRAMS THAT ARE ENDURING.

It is ideal that JPE programs be created by constitutional provision or statute. Statutory authorization provides a balanced approach, ensuring some degree of permanence coupled with the flexibility to adopt improvements as needed. Where constitutional or statutory authorization is not practical, programs can also be established by court rule, or even a partnership with the bar and civic organizations.

5.2

PROVIDE PROGRAMS WITH THE RESOURCES NEEDED TO EFFECTUATE THE GOALS OF JPE.

JPE programs require ongoing, adequate, and secure funding. This is especially important given the sophisticated needs programs have today, the additional data and analysis required, modern approaches to evaluations, and a commitment to performance improvement.

When JPE programs face insufficient funding, they are undervalued and understaffed. Understaffing is a major factor that limits the effectiveness of JPE programs. Programs need staff to have the capacity to elevate and modernize. We hope that each JPE program will take up these recommendations, review and evaluate their own programs, and implement improvements. Doing so requires sufficient capacity, and this should be supported by leadership and by investment in staffing.

Funding needs will vary by state, but aside from investment in staffing, some examples of work that should be prioritized for funding include consultation with survey research experts, a clear and strategic communications plan with supporting components (i.e., a website), and outreach to inform voters about evaluation programs. In addition to funding JPE, it is important that states fund professional development programs in connection with JPE, likely through the office of judicial education, to enhance capacity to support the performance improvement of judges.

5.3

SEEK BUY-IN AND COOPERATION FROM ALL STAKEHOLDERS.

Symbolic and substantive commitment to the program from the judiciary, legislators, and the bar is essential. Court leaders need to speak publicly about the value of JPE to the judicial branch and help judges use evaluation results to improve their performance. There also needs to be

cooperation with courts on data and professional development. Individual judges should encourage those who appear before them to provide feedback and should be open to this feedback. Legislators should raise awareness of JPE programs among constituents and ensure that such programs are funded. The bar should remind attorneys to participate in evaluating judges, whether their feedback is positive, neutral, or negative. Members of the public should consider serving on a commission. They should also take steps to inform themselves as voters and encourage friends and family to do the same.

RECOMMENDATION 6

ENSURE AND IMPROVE TRANSPARENCY FOR INTENDED AUDIENCES.

6.1 Be clear with the public about how JPE works and what the evaluations mean.

6.2 Ensure judges have access to accurate information about the JPE process, including reasons for the commission's findings.

There is a delicate balance between transparency and performance improvement, and the right balance for each state depends on several factors, including the purpose and structure of its JPE program. Some JPE methods that promote transparency and public accountability (such as publishing certain details on evaluations) might disincentivize performance improvement by making judges feel defensive and thus less likely to use the feedback. Private feedback creates a safer environment for reflection and receptivity to feedback but provides less transparency and public accountability. The recommendations about transparency that follow need to be considered in light of this tension.

6.1

BE CLEAR WITH THE PUBLIC ABOUT HOW JPE WORKS AND WHAT THE EVALUATIONS MEAN.

JPE programs typically find that judges meet performance standards. This can lead members of the public to conclude that JPE is simply a “rubber stamp” process. However, a merit selection system is designed to select highly qualified judges from the outset, so positive evaluations are expected. Also, judges are usually alerted to a pending “does not meet performance standards” recommendation and choose to resign rather than stand for reelection, so their reports never reach the public.

Even when the process is working, that may not be clear to the public. Funding should be dedicated to publication and promotion of the JPE results and education about the process. Evaluation results should be widely disseminated to the public.⁴⁰ The vote count on the commission’s assessment of individual judges’ fitness for performance should be publicly available. In states where judicial performance is evaluated for self-improvement only and/or where judges are not subject to reselection, programs should consider the positive impact on public trust and confidence in the judiciary of providing at least aggregate results for each level of court. To mitigate the concern about rubber stamping, states can report aggregate judicial resignation numbers to communicate the complete picture of JPE’s effectiveness without divulging personal information.⁴¹

6.2

ENSURE JUDGES HAVE ACCESS TO ACCURATE INFORMATION ABOUT THE JPE PROCESS, INCLUDING REASONS FOR THE COMMISSION’S FINDINGS.

In the convenings, judges shared that more information about the process would improve their overall trust and comfort with JPE. When judges do

not trust the process, they are less likely to change their behavior based on JPE feedback.

JPE administrators should provide judges clear information about how their evaluations will work and what they will entail. This can include information related to specific evaluations, as well as training for judges about the process generally.

Once the evaluation is conducted, judges should have the opportunity to review all the information the commission considers. In addition, JPE programs need to structure evaluations for judges to contextualize the negative comments. Judges shared that while they welcome feedback, it can be emotionally difficult and unhelpful to receive harsh critiques without any context or support. Rather than providing judges with evaluations without context or additional support, a supportive structure—such as training or mentoring—can be beneficial. This structure can provide perspective, normalize critiques as an inherent part of public service, and help judges interpret feedback. JPE programs can also help by organizing feedback in a way that is easier to digest, identifying comments as positive, negative, or neutral. JPE programs should also give judges the chance to review their assessments and challenge aspects they deem inaccurate or unfair before they are made public. If the judge’s concerns are not considered in revising the report, the judge should be permitted to make a statement in or adjacent to the public report.

Some states involve judges in the process by soliciting feedback after JPE evaluations have been completed or by inviting judges to sit on committees tasked with improving JPE. In general, improved communication with judges about the process can elevate confidence.

CALL TO ACTION: WHAT COMES NEXT?

This report offers a comprehensive framework for modernizing JPE, but its value lies in its application. To translate these recommendations into tangible improvements, states and courts must take concrete action. This section outlines next steps for embracing this call to action and implementing these recommendations, both for states with existing JPE programs and those without.

Next Steps for States *with* JPE Programs

These recommendations set forth a vision for JPE programs that all stakeholders can trust is rigorous but fair, with clear and objective performance standards, reliable data reflecting a range of perspectives, transparency to the public, and meaningful support for judges to implement feedback. States can take the following steps to bring about that vision for their programs:

- ▶ **Review and Share the Recommendations.** Review these recommendations and the principles they are built on. Share them with relevant stakeholders, leaders, and decision-makers in your state to help them understand the importance of modernizing JPE and the benefits of implementing these practices.
- ▶ **Assess Your Program.** Conduct a review of your JPE program, comparing it to the best practices outlined in this report. Identify your program's areas of strength as well as opportunities for improvement.
- ▶ **Define Challenges and Needs.** Each state has unique features, needs, and circumstances and should adapt these recommendations accordingly. Identify your program's challenges and barriers, understand their impact, and assess potential opportunities for improvement.
- ▶ **Identify a Core Team.** Strong leadership is necessary to bring about reform. Identify a core team or working group of stakeholders who can lead implementation efforts. This team will determine the scope and parameters of program changes at the outset of the process. In creating this team, consider balance and diversity and ensure key perspectives are represented.
- ▶ **Engage Stakeholders.** Effective JPE programs require buy-in from all stakeholders, including anyone with decision-making authority over JPE in your jurisdiction. Possible stakeholders include judges, JPE commissioners, legislators, court administrators, clerks, lawyers, court users, and members of the public. States should work with stakeholders to seek feedback, discuss these recommendations, build internal and external support, and assess next steps.
- ▶ **Create a Vision and Goals.** Based on your assessment and stakeholder input, create a clear vision for the future of your JPE program. Establish specific, measurable, and achievable goals for bringing about this vision, as well as clear timelines for implementation.
- ▶ **Develop Tailored Recommendations.** Tailor the recommendations in this report to the needs of your program, prioritizing those that will have the greatest impact. Develop a plan for implementation of tailored recommendations, considering what resources you need to address challenges and align your program with best practices.
- ▶ **Take Action.** Coordinate with relevant stakeholders to execute the plan for implementation. Invest in communication and education about improvements and changes to your program. Once recommendations have been implemented, be prepared to continually iterate and adjust your approach based on experience and feedback.

- ▶ **Share and Collaborate.** Share your experiences and lessons learned with other states to foster a dialogue on JPE best practices and opportunities.

Courts are always changing, as are the needs and expectations of users and judges alike. To remain effective and maintain its integrity, JPE must continually modernize. By taking a strategic approach to implementing these recommendations, JPE programs can enhance their programs to better effectuate the goals of JPE.

Next Steps for States *without* JPE Programs

Many of these recommendations are designed to enhance existing JPE programs. However, the principles underlying these recommendations have implications for all courts, beyond the context of formal JPE programs. They emphasize the evolving role of judges, the importance of public trust, and the need for ongoing professional development. Embracing the principles underlying these recommendations can strengthen public trust and judicial performance, even in the absence of a formal JPE program. For states without established JPE programs, these recommendations offer a framework that states can apply in the following ways:

- ▶ **Embrace the Principles.** While a formal JPE program may not be immediately possible, embrace the core principles of fairness, transparency, and continuous improvement in all aspects of your court system.
- ▶ **Define Clear Performance Metrics.** Develop clear, objective, and measurable criteria for judicial performance that reflect the role of a judge and focus on the process of judging (not on the outcomes of specific cases). Such criteria set clear expectations for judges and court users and create a foundation for accountability, even absent a structured evaluation process. They can be used for judicial selection, retention, and professional development.
- ▶ **Invest in Professional Development for Judges.** Create opportunities for professional development to support judges in enhancing their skills and addressing areas for growth. Professional development opportunities can include, for example, feedback, self-reflection, mentorship programs, and training. By promoting a culture of continuous improvement, states can bolster public trust in courts.
- ▶ **Foster Open Communication and Transparency.** Promote open communication between the courts and the public. This is important for building trust and can include, for example, proactively sharing information about court operations and seeking feedback from court users.
- ▶ **Engage with Stakeholders.** Engage in conversations with relevant stakeholders about the need to embrace core JPE principles, the potential benefits of a JPE program, and opportunities for future development. Possible stakeholders can include judges, court administrators, bar associations, legislators, and the public.
- ▶ **Explore Resources.** Use resources like this report and other IAALS tools on JPE to guide your efforts in building a more accountable and transparent judicial system.

Even without a formalized evaluation structure, states can promote enhanced judicial performance and public confidence in the judiciary by embracing the best practices underlying these recommendations. These best practices demonstrate a commitment to accountability, transparency, and support of judges, laying the groundwork for the possibility of developing a more formal JPE process in the future.

CONCLUSION

As courts across the country grapple with issues of bias, transparency, accountability, and public trust, these recommendations provide a path forward for modernizing judicial practices to support the judiciary's position as a trusted institution. IAALS remains committed to advancing effective judicial performance evaluation and ensuring every state has access to the tools and knowledge to build a judicial system that earns and maintains the public's trust.

ENDNOTES

- 1 This report uses the title "judicial performance evaluation." However, it should be noted that different states use different titles (e.g., judicial performance review).
- 2 Jordan M. Singer, IAALS, *Judicial Performance Evaluation in the States: The IAALS JPE 2.0 Pre-Convening White Paper 1-2* (2022).
- 3 *Id.* at 2.
- 4 David C. Brody, *The Use of Judicial Performance Evaluation to Enhance Judicial Accountability, Judicial Independence, and Public Trust*, 86 *Denv. U. L. Rev.* 115, 131-145 (2008).
- 5 *Id.* at 156-57.
- 6 Singer, *supra* note 2, at 5.
- 7 See generally *id.*; Natalie Knowlton & Malia Reddick, IAALS, *Leveling the Playing Field: Gender, Ethnicity, and Judicial Performance Evaluation* (2012); Rebecca Love Kourlis & Malia Reddick, IAALS, *Transparent Courthouse Revisited: An Updated Blueprint for Judicial Performance Evaluation* (2016); Rebecca Love Kourlis, Malia Reddick, & Natalie Anne Knowlton, IAALS, *Recommended Tools for Evaluating Appellate Judges* (2013); Rebecca Love Kourlis & Jordan M. Singer, *Using Judicial Performance Evaluations to Promote Judicial Accountability*, 90 *Judicature* 201, 206 (2007).
- 8 Danielle Kalil et al., IAALS, *JPE 2.0: National Perspectives on Judicial Performance Evaluation*, at 12 (2024) [hereinafter *JPE Perspectives Survey*].
- 9 *JPE Perspectives Survey*, *supra* note 8, at 12.
- 10 *Id.* at 8.
- 11 *Id.* at 17.
- 12 *Id.* at 18.
- 13 *Id.* at 8.
- 14 See, e.g., CBA-CJI Diversity on the Bench Coalition, *2024 Report and Recommendations: Retention (2024)* (finding that preparing for retention was a main stressor among judicial officers in Colorado).
- 15 *JPE Perspectives Survey*, *supra* note 8, at 18.
- 16 *JPE Perspectives Survey*, *supra* note 8, at 19.
- 17 *Id.*
- 18 For more on this topic, see Rebecca Love Kourlis & Malia Reddick, IAALS, *Judges Aren't Sexy: Engaging and Educating Voters in a Crowded World* (2016).
- 19 See the Colorado Office of Judicial Performance Evaluation webpage for examples.

- 20 The National Center for State Court's CourTools Trial Court Performance Measures offer clear and measurable criteria for assessing the performance of trial courts, including objective data such as clearance rates, time to disposition, and age of active pending caseload.
- 21 For sample survey questions, see *Leveling the Playing Field*, *supra* note 7, at Appendix B and C.
- 22 Not every survey respondent will be able to assess each criteria. For example, a court user may not have information about a judge's administrative capacity or efforts to advance the justice system. Strategies for navigating this are discussed in Recommendation 3.1.
- 23 American Bar Association, *Black Letter Guidelines for the Evaluation of Judicial Performance* (2005).
- 24 For sample survey questions, see *Leveling the Playing Field*, *supra* note 7, at Appendix B and C.
- 25 C.R.S. § 13-5.5-107(f) (2024).
- 26 *Transparent Courthouse Revisited*, *supra* note 7, at 8 (finding biases were more pronounced in states that used attorney-only surveys).
- 27 *JPE Perspectives Survey*, *supra* note 8, at 11-13.
- 28 *Transparent Courthouse Revisited*, *supra* note 7, at 10-11.
- 29 *JPE Perspectives Survey*, *supra* note 8, at 12.
- 30 *JPE Perspectives Survey*, *supra* note 8, at 8, 12, 16.
- 31 See generally NCSC, *CourTools Trial Court Performance Measures*, *supra* note 20. This tool offers practical guidance for measuring objective case management data across a number of metrics, including clearance rates, time to disposition, age of active pending caseload, trial date certainty, reliability and integrity of case files, legal financial obligations, effective use of jurors, court employee satisfaction, and cost per case.
- 32 *JPE Perspectives Survey*, *supra* note 8, at 13-14.
- 33 For examples from Utah, see generally Woolf, Nicholas H. and Yim, Jennifer MJ, *The Courtroom-Observation Program of the Utah Judicial Performance Evaluation Commission*, *Court Rev.: The Journal of the American Judges Association*, 368 (2011) (describing Utah's observation procedures, including recruitment and training of volunteer observers).
- 34 Alaska is an example of a state that provides reversal rates as well as the comparative context for those rates. See Alaska Judicial Council, *Memorandum on Appellate Evaluation of Judges Eligible for Retention in 2020* (August 7, 2020).
- 35 *JPE Perspectives Survey*, *supra* note 8, at 12.
- 36 See NCSC, *CourTools Trial Court Performance Measures*, *supra* note 20. The Measure 1: Access and Fairness tool offers examples and practical guidance for capturing "ratings of court users on the court's accessibility and its treatment of customers in terms of fairness, equality, and respect."
- 37 *JPE Perspectives Survey*, *supra* note 8, at 5.
- 38 *JPE Perspectives Survey*, *supra* note 8, at 17-18.
- 39 See generally David Swenson, Ph.D. L.P. et al., American Bar Association, *Stress and Resiliency in the U.S. Judiciary, 2020 Journal of the Professional Lawyer* (2020) (underscoring the need for wellness resources for judges, for whom "the daily challenges of the position are stressful and have an adverse effect on a judge's health and well-being"). Colorado's work on judicial well-being offers one example of such resources, including wellness tools, a committee responsible for making recommendations to the judicial branch, a Judicial Wellness Roundtable, and various groups and discussion forums.
- 40 For more on this topic, see generally *Judges Aren't Sexy: Engaging and Educating Voters in a Crowded World*, *supra* note 18.
- 41 Colorado offers one example of how to do this. See Colorado Office of Judicial Performance Evaluation, *2024 Judicial Performance Evaluation Statical Report* (2024) ("[116] judicial officers declared their intent to stand for retention and appeared on the ballot in 2024. . . . Commissions evaluated four additional judges not standing for retention.")



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

How can we help? Consulting Inquiries: iaals@du.edu