Published by the Bolch Judicial Institute at Duke Law. Reprinted with permission.
© 2026 Duke University School of Law. All rights reserved. JUDICATURE.DUKE.EDU

JPE 2.0: A MODERN APPROACH TO EVALUATING JUDICIAL PERFORMANCE

Judicial performance evaluation programs were designed to improve judging and inform the public. The authors argue that modernizing them can strengthen both judicial performance and public trust.

By Jordan M. Singer & Danielle Kalil

Judicial performance evaluation (JPE) programs have long been used to provide feedback to state judges and the public about judges' performance on the bench. Courts can utilize JPE to advance the professional development of judges through individualized reports, training, and mentoring programs. Voters and officials can review JPE results for information relevant to judicial retention or reappointment decisions. And because JPE emphasizes the process of judging rather than discrete case outcomes, it can help direct public discourse away from partisan characterizations of the courts and toward apolitical judicial qualities like impartiality, clear communication, and professional demeanor.[1]

Spurred by the promise of these benefits, JPE programs proliferated rapidly between the early 1980s and the mid-2000s.[2] By 2010, formal JPE programs were operating in nearly half the states. Over the past decade, however, enthusiasm for JPE programs has waned. Some critics have expressed concern about the accuracy and validity of survey data. Others, almost exclusively from outside the judiciary, have complained that JPE programs fail to account explicitly for judicial discipline or judicial ideology. Still others have urged JPE programs to become more transparent. A small number of states have discontinued or defunded their JPE programs.[3]

The Institute for the Advancement of the American Legal System (IAALS) at the University of Denver is a longtime leader in JPE research and advocacy. It resolved to examine and respond to these trends. Over nearly four years, IAALS held conversations and convenings with JPE stakeholders, analyzed the history and critiques of JPE, and solicited ideas and

Published by the Bolch Judicial Institute at Duke Law. Reprinted with permission.
© 2026 Duke University School of Law. All rights reserved. JUDICATURE.DUKE.EDU

Published by the Bolch Judicial Institute at Duke Law. Reprinted with permission.
© 2026 Duke University School of Law. All rights reserved. JUDICATURE.DUKE.EDU

creative solutions. The clear takeaway from these conversations was that JPE remains fundamentally sound but should be modernized to provide more reliable data, incorporate the views of more court users, and better reflect the roles and responsibilities of the modern judiciary.

This article discusses the history of JPE, the JPE 2.0 project, and IAALS recommendations for modernizing JPE.

EARLY DEVELOPMENT OF STATE JPE PROGRAMS

The first formal JPE program was authorized in Alaska in 1975, but the practice of periodically "judging the judges" is much older.[4] In the 19th century, as judicial elections became more widespread, local bar associations began polling their members about judicial candidates and releasing the results before the election as a rather crude endorsement (or nonendorsement) tool.[5] Bar polls eventually developed into more detailed examinations of each candidate's judicial qualities, but even these later polls had significant methodological flaws. Many polls, for example, did not require attorneys to have appeared before the judges they were evaluating.[6] Given the dearth of other information about judicial performance, however, the polls gained traction. By the late 1970s, they were widespread throughout the country.[7]

In 1980, the American Bar Association (ABA) created a task force to examine the use of judicial evaluations for a different purpose: professional development and improvement.[8] The ABA recognized that bar polls alone would not provide information adequate to help judges develop their professional skills and consequently pushed for programs that would take additional, objective information (such as docket management statistics) into account.

Because JPE emphasizes the process of judging rather than discrete case outcomes, it can help direct public discourse away from partisan characterizations of the courts.

The ABA also recommended expanding evaluations to include information from court users other than attorneys, as well as creating independent committees with lay representatives to conduct the evaluations. In 1985, the ABA officially approved its Guidelines for the Evaluation of Judicial Performance.[9] With support from the ABA and other prominent national organizations, several states — including Colorado, Connecticut, and New Jersey — initiated JPE programs in the 1980s.[10]

Formal JPE programs grew steadily in the 1990s and 2000s, encompassing states with a variety of judicial-selection mechanisms. JPE was particularly popular in merit-selection states, where it was championed as a way to get reliable information about judicial performance to voters in advance of retention elections. But JPE was also implemented in states with contested judicial elections, states in which judges were selected by the legislature or a commission, and even states (and one federal court) in which judges were effectively appointed for life.[11] In these jurisdictions, JPE was seen as a valuable professional development tool.

The proliferating JPE programs had somewhat different structures and were authorized in different ways: by constitutional amendment, by legislation, by court rule, or by informal partnerships with the bar.[12] What they had in common was a commitment to providing meaningful feedback to judges and, when appropriate, the public. The success of JPE programs was such that in 2005, when the ABA revisited its recommendations from 20 years earlier, it readily recommitted itself to JPE and its underlying principles.[13]

A BUMPY ROAD IN THE 2010s

At the high-water mark of judicial performance evaluation in 2010, 19 states, as well as the District of Columbia and Puerto Rico, had instituted official JPE programs.[14] During the next decade, however, only two new programs were created, and four programs (including one that had started just two years earlier in North Carolina) were effectively shuttered.[15] At the same time, critiques of existing programs began to surface.

The most substantial criticisms took aim at the surveys distributed to attorneys, court staff, and others like jurors and litigants. Surveys have long been at the core of JPE programs and are arguably the best method for capturing court users' perceptions of judicial behavior and due process.[16] Most state survey instruments, however, had not been updated in many years, and some studies suggested that the framing of survey questions led to disproportionately low scores for women and judges of color.[17] Attorney surveys in particular were also criticized for low response rates and for occasionally eliciting unprofessional or inappropriate comments about a judge under the cloak of respondent anonymity.[18]

Some people also demanded greater transparency and communication

Published by the Bolch Judicial Institute at Duke Law. Reprinted with permission.
© 2026 Duke University School of Law. All rights reserved. JUDICATURE.DUKE.EDU

about the JPE process. In states where evaluations were shared only with the evaluated judge, survey respondents worried about how their feedback was used, or whether it was used at all.[19]

In states that made evaluations public, some observers expressed concern that the JPE process and purpose behind it were not adequately explained to the public or that the results were inadequately publicized.[20]

A particularly beguiling issue involved the hazy relationship between judicial evaluation and judicial discipline. Traditionally these were separate matters that were handled by separate state boards, and with good reason: Judicial discipline mechanisms were designed to investigate and address impropriety on the bench, while JPE programs were designed to investigate and address judicial skill and competence. Because these programs served separate functions and judicial discipline is often addressed privately, judicial discipline records were typically not part of JPE evaluations. Our convenings found that a growing sector of the public began to conflate the two matters, arguing that if a judge was under investigation or had been disciplined, it should be reflected in the ordinary evaluation process. The firewall between discipline and JPE, while complex and historically justified, became increasingly difficult to explain.

THE JPE 2.0 PROJECT

These critiques, as well as the withdrawal of official support for JPE programs in Florida, Kansas, North Carolina, and Tennessee, drew the attention of IAALS's JPE Working Group, a national consortium of JPE program coordinators and researchers.[21] In 2021, IAALS convened a subset of Working Group members to create the JPE 2.0 Task Force to investigate and respond to these concerns — and renew JPE as a valued and trusted process. The task force recognized that JPE had an important historical pedigree and also appreciated that it must remain responsive to contemporary needs.[22]

The task force resolved to collect perspectives from a wide range of JPE stakeholders — judges, lawyers, court users, government officials, and the general public — and to bring them together to discuss if, and how, JPE programs could better serve them moving forward.[23] To this end, the task force undertook three initiatives: a survey on judges' perspectives on JPE, a 2022 convening of a broad range of stakeholders to address key challenges, and subsequent virtual convenings of specific stakeholder groups in 2023.[24]

JPE PERSPECTIVES SURVEY

The task force's first order of business was to ascertain how sitting judges felt about the evaluation process. Although judges are obviously affected by the structure and substance of JPE programs, to that point, little effort had been made to collect their views comprehensively and systematically.[25] In late 2021 and early 2022, IAALS administered the JPE Perspectives Survey to judges in eight states with JPE programs. Nearly 700 judges responded, providing thousands of quantitative data points and hundreds of open-ended comments.[26]

The survey offered an important glimpse into the minds of those most directly affected by performance evaluations: the judges themselves. The findings were mixed. On one hand, most respondents had a positive view of the JPE process in their home states, with 68.1% of respondents agreeing or strongly agreeing that "Overall, I am satisfied with the JPE/JPR process," 72.1% agreeing that "Going through ▸

Published by the Bolch Judicial Institute at Duke Law. Reprinted with permission.
© 2026 Duke University School of Law. All rights reserved. JUDICATURE.DUKE.EDU

the JPE/JPR process has been beneficial to my professional development," and nearly 82% indicating that they felt adequately informed about their state's JPE process.[27]

On the other hand, respondents expressed less confidence that JPE helped their relationship with the general public. Only about 26% agreed that their JPE program increased their judicial independence, and only about 33% agreed that it "helps the public understand the work that I do."[28] Interestingly, almost 37% of respondents indicated that JPE "does not increase my accountability to the public."[29]

The survey also asked, "When you think about JPE/JPR in your state as a whole, do you have specific concerns about the evaluation process?" More than 58% of judges responded affirmatively.[30] The survey invited respondents to expand upon their answers through open-ended comments. The comments were often state-specific, but three broader themes emerged.

First, while appreciating attorneys' concerns for anonymity, many judges stated that receiving anonymous feedback made it difficult for them to improve because they were unable to put it in context.

Second, several judges expressed concern about low response rates for attorney surveys, especially because they believed that attorneys with poor impressions of the judge would be more motivated to complete the survey, thereby skewing the overall data.

Third, many comments indicated a high level of judicial stress and fear about JPE, with an associated belief that the process was being weaponized by legislatures and lawyers.[31]

JPE STAKEHOLDERS CONFERENCE

While the JPE Perspectives Survey was being administered, the JPE 2.0 Task Force began a close examination of the criteria used to evaluate judges.[32] Every state with a formal JPE program has relied on the five categories articulated by the ABA in 1985: legal knowledge, impartiality, clarity of communication, professional demeanor, and administrative capacity. The ABA itself borrowed these criteria from earlier bar polls, concluding that there was "substantial agreement upon many of the standards by which judges could be evaluated."[33] The task force concluded that these criteria still made sense but nevertheless felt incomplete in light of substantial changes to judges' day-to-day roles since the 1980s.[34] Among other things, it noted that judges today confront more robust case-management responsibilities, more frequent interaction with self-represented litigants, and far more advanced technology than their predecessors 40 years ago.[35] Perhaps revisiting the criteria was in order.

IAALS next brought together a broad cross-section of JPE stakeholders at a national convening in May 2022. The approximately two dozen participants included representatives from the judiciary, the bar, court administration, state JPE commissions, state executive and legislative branches, the legal academy, and advocacy groups for court users. Two days of robust discussion and debate ensued. Participants generally agreed that JPE programs continued to serve a valuable function, both in educating the public about the role and activities of the judiciary and in educating judges about their individual and collective strengths and weaknesses. At the same time, participants identified several trends and practices that were hampering JPE's growth and acceptance.

The first set of issues concerned technical challenges to collecting complete and reliable information on judicial

Published by the Bolch Judicial Institute at Duke Law. Reprinted with permission.
© 2026 Duke University School of Law. All rights reserved. JUDICATURE.DUKE.EDU

performance. Participants agreed that data reliability could be improved by universally adopting evidence-based practices for crafting and administering surveys. They also recommended using modern courthouse technology to generate additional data that could supplement and contextualize survey responses. For example, modern docket-management software can provide a clear overview of the types of cases and litigants who have appeared before the judge, and videoconferencing software can allow for unobtrusive courtroom observation.

A second set of issues related to trust in JPE programs. A recurring theme among all stakeholder groups was suspicion or skepticism about how the results might be used. Attorneys feared that they would be personally identified by their survey responses and subject to judicial reprisal. Judges felt that attorneys misused surveys to vent hurtful and unconstructive comments and that JPE commissioners did not do enough to provide constructive feedback. More broadly, stakeholders noted that longitudinal studies have shown a more general decline in public trust in government and institutions,[36] which might mean that the public would look upon JPE results with skepticism, even if JPE were perfectly executed.

This is not to say that JPE programs themselves are untrustworthy, or even that most stakeholders at the 2022 convening did not trust them. But participants felt that the level of trust across the board could be improved. Lack of trust in the JPE process also has ancillary consequences. When public confidence in JPE drops, it becomes easier for political actors to weaponize specific JPE results for partisan gain (e.g., by targeting a judge in a retention election). Diminished trust can also contribute to inadequate institutional support for JPE. If stakeholders feel JPE is not worthy of their confidence, they are more likely to withhold adequate funding from JPE programs or decline to participate in the evaluation process.

JPE remains fundamentally sound but should be modernized to provide more reliable data, incorporate the views of more court users, and better reflect the roles and responsibilities of the modern judiciary.

The third and perhaps most daunting set of issues concerned conceptual hurdles to JPE's success. Consistent with the initial reflections of the JPE 2.0 Task Force, participants affirmed that the traditional evaluation criteria were meaningful but did not fully capture the modern judge's roles and responsibilities. For example, they generally agreed that judges in the 2020s should be able to demonstrate competence with court technology and work effectively with self-represented litigants, but noted that neither skill was directly evaluated by a typical JPE program. Similarly, participants discussed that a modern understanding of the justice system would account not only for the views of attorneys and court staff but also for the perspectives and experiences of litigants, jurors, and other court users. If JPE were to meaningfully assist in professional development and promote public understanding of the courts, it would need to reflect these new judicial roles and capture the perspectives of all stakeholders.

Importantly, participants agreed that none of these challenges was insurmountable. Rather, each represented an opportunity to improve the quality of, and confidence in, judicial evaluations. Participants stressed that JPE programs could do more to avoid errors in data collection. Mutual trust and respect between judges, court users, and the public could be strengthened. Criteria could be modernized to better reflect the current judicial role. The critical question was how to achieve these goals while gaining buy-in from different groups of JPE stakeholders.

VIRTUAL CONVENINGS

In 2023, IAALS followed the national convening with four virtual convenings.[37] Each gathered representatives of a specific stakeholder group — judges, JPE commissioners, attorneys, and all other stakeholders — to delve into the challenges and opportunities posed by existing JPE programs. The virtual convenings included a mix of alumni from the national convening and new participants from around the country.

Building on primary themes from the national convening, the virtual ones featured moderated discussions around three key sets of questions:

- In what aspects of the JPE process do you have the most and the least confidence?
- What concrete steps could be taken to increase your confidence (and public confidence more generally) in JPE programs?
- Do existing JPE criteria adequately capture the roles and responsibilities of the modern judge? Should criteria be added or modified? ▸

Published by the Bolch Judicial Institute at Duke Law. Reprinted with permission.
© 2026 Duke University School of Law. All rights reserved. JUDICATURE.DUKE.EDU

Across all four virtual convenings, participants expressed the most confidence in what might be termed "interpersonal trust." For example, both judges and attorneys showed confidence in the experience and sincerity of JPE commissioners. Commissioners, too, praised opportunities to observe and interact with judges directly. All groups felt that despite their flaws (addressed in the next paragraph), surveys were an important starting point for evaluation as they allowed for widespread participation in the evaluation process and could help identify themes to begin conversations with judges.

Participants in the virtual convenings expressed the least confidence in surveys and survey administration. Their concerns echoed those raised in the national convening as well as in the JPE Perspectives Survey responses: low response rates, the risk of bias or inappropriate comments, and the failure to capture nonattorneys' views. Attorney participants echoed concerns that their responses were not truly confidential.

Participants suggested a number of concrete steps to improve both survey quality and response rates, including offering incentives (such as continuing legal education credit), providing more information about how survey responses are used and how feedback is kept confidential, crafting survey instruments to reduce the risk of bias, shortening surveys, and diversifying the methods of evaluation beyond surveys to provide a more complete view of a judge's performance. Across the virtual convenings, there was widespread consensus that the traditional categories of evaluation criteria should be maintained but that the definitions of and approaches to those criteria needed to evolve. In addition, several participants suggested adding new criteria related to technological competence and judicial innovation. Participants noted that judicial temperament is the most challenging criterion to measure objectively but broadly agreed it should be included because it is important that judges show respect for court users on and off the bench. Some participants pointed out that recordings of courtroom proceedings and the use of questions about the judge's specific behavior could add objectivity to the analysis. Judges stressed that evaluation criteria should reflect the tasks, roles, and responsibilities that they are called on to perform in their respective courts.

> Education and outreach that explain both the existence and purpose of JPE can help prevent misconceptions about the judicial role and build increased confidence in the judiciary.

The findings from the national and follow-up virtual convenings made clear that JPE programs are still worthy endeavors. To reach their full promise, however, they should better reflect the contemporary roles and responsibilities of judges as well as the needs and expectations of court users.

A MODERN APPROACH TO JUDICIAL PERFORMANCE EVALUATIONS

The JPE 2.0 Task Force met in late 2023 to analyze the findings and themes of the JPE Perspectives Survey and the convenings. Building on the task force's work, IAALS released a report and key recommendations for modernizing JPE programs in May 2025:[38]

- Foster trust and confidence in the JPE process
- Modernize evaluation criteria to account for the changing roles and responsibilities of the judiciary
- Implement measures to improve information and reduce errors in evaluations
- Build a culture of and embrace judicial performance improvement as a key aspect of JPE
- Display institutional commitment to JPE
- Ensure and improve transparency for intended audiences

Foster trust and confidence in the JPE process. A key takeaway from the convenings and JPE Perspectives Survey is that JPE only works if stakeholders have confidence in the integrity of the process. JPE programs can earn that confidence by embodying the same values that the public expects of judges, including impartiality, clear communication, and administrative efficiency.

Impartiality is communicated in part through effective design. JPE programs should ideally be housed in an office independent of the judiciary or any other branch of government. Evaluation commissions should be structured to invite a broad array of knowledge and perspectives into the evaluation process. JPE commissioners should receive formal training on their roles and responsibilities, and commissions should adhere to well-articulated procedures.

Communication about the JPE process should be proactive and seek to engage and educate the public about why JPE is so important. In states where JPE is

Published by the Bolch Judicial Institute at Duke Law. Reprinted with permission.
© 2026 Duke University School of Law. All rights reserved. JUDICATURE.DUKE.EDU

used for retention elections, outreach to the public (and the media) about where to find evaluation results and how to interpret them can help create a more informed voter base. In all states, education and outreach that explain both the existence and purpose of JPE can help prevent misconceptions about the judicial role, build increased confidence in the judiciary, and create more informed stakeholders.

Administrative efficiency can be improved by coordinating the timely exchange of information between JPE and judicial discipline programs. While these programs have different purposes and should remain separate, public trust can diminish if judicial evaluations fail to account for even well-publicized disciplinary actions. Articulating circumstances in which JPE and disciplinary proceedings might cooperate — for example, when public disciplinary proceedings are happening simultaneously to a performance evaluation — can boost public confidence in both programs and in the judiciary as a whole.

Modernize evaluation criteria to account for the changing roles and responsibilities of the judiciary. A recurring theme of the convenings was that judging has changed since the ABA first endorsed JPE criteria in the mid-1980s. Judges today must possess proficiency with courtroom technology, hands-on case management, and the needs of a large and increasing number of self-represented litigants. These changes have shifted how judges both engage with parties and run their chambers and courtrooms.

In light of these developments, JPE 2.0 offers revisions to the scope and application of the traditional evaluation criteria, as well as the addition of a new, sixth criterion:

- Judicial competence (replacing "legal knowledge")
- Impartiality
- Clarity of communication (replacing "clarity of written and oral communication")
- Judicial temperament
- Administrative capacity
- Efforts to advance the justice system (new criterion)

The first criterion for judicial performance is traditionally called legal knowledge, meaning knowledge of substantive law and of procedural and evidentiary rules, and demonstrated legal reasoning ability. As a benchmark for judicial performance, it is thoughtful but incomplete.

A better and more comprehensive criterion is judicial competence, defined as not only a judge's legal knowledge but also competence with technology, knowledge of the community, and ability to apply the relevant law, tools, and practices in a manner appropriate to the specific needs of the parties and the case. Importantly, judicial competence may look different for different judges; for example, new judges often face a steep learning curve.

The second traditional criterion, impartiality, remains highly relevant to a judge's performance. JPE 2.0 maintains the classic standards for impartiality (such as treating people with dignity and respect and avoiding impropriety or the appearance of impropriety). Consistent with fundamental due process principles, judges should also be reviewed for equal, unbiased treatment and fair and impartial application of the law.

The third criterion, traditionally known as the clarity of written and oral communication, has been replaced by the more general term "clarity of communication." This principle recognizes ▸

Published by the Bolch Judicial Institute at Duke Law. Reprinted with permission.
© 2026 Duke University School of Law. All rights reserved. JUDICATURE.DUKE.EDU

that all forms of communication that emanate from a judge's courtroom or chambers should be straightforward and understandable to all relevant audiences. Judges should adapt their communication styles to the needs of the audience, favoring plain language where appropriate. Moreover, judges should be responsible for the clarity and tone of all communications coming from their respective chambers, even if not written by them personally.

The fourth traditional criterion, judicial temperament, is intended to capture observable judicial behavior that influences trust and confidence in the justice system. It includes acting with patience and self-control, treating people with appropriate courtesy, maintaining control over court proceedings, displaying attentiveness to the proceedings and parties, and exhibiting sensitivity to cultural differences. Because judicial temperament can mean different things to different people, JPE programs should provide clear definitions and concrete examples of specific judicial behaviors to survey respondents and courtroom observers. JPE programs should also recognize that the appropriateness of judicial behavior in any given instance may depend on context.

The fifth traditional criterion, administrative capacity, addresses each judge's ability to resolve cases promptly and fairly. It encompasses time management, proactive case management, effective use of technology, and effective use of innovative tools or processes. JPE 2.0 explicitly adds the caveat that the assessment of a judge's administrative capacity should be aligned with that judge's particular roles and responsibilities, as well as the characteristics of the cases and parties coming before the judge. Importantly, this criterion should not encompass any aspects of administration outside the judge's direct control, such as clerk's office procedures or mandates from the state administrative office or state supreme court.

The final criterion, the judge's efforts to advance and improve the justice system, is new. While a minority of JPE states evaluate a judge's involvement in the community, none specifically seek to capture the judge's efforts to improve the legal system, educate the public about the work of the courts, or advance access to justice. The new criterion accordingly includes the judge's contributions to the legal community as well as the local community, and contemplates a wide range of activities inside and outside the court system. Mentoring, committee work, public education and outreach, legal system innovation, and other efforts to improve the system or existing practices should all be taken into account. The goal of including this new criterion is to convey that 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.

Implement measures to improve information and reduce errors in evaluations. The JPE Perspectives Survey and JPE 2.0 convenings highlighted opportunities to modernize the collection of information on judicial performance. The IAALS recommendations reflect the significant technological and methodological advances in survey administration since JPE was first developed, as well as new insights on gathering user perspectives in our justice system.

First and foremost, surveys should use evidence-based best practices to increase fairness and reliability. This begins with a broad respondent base that includes not only attorneys but

Published by the Bolch Judicial Institute at Duke Law. Reprinted with permission.
© 2026 Duke University School of Law. All rights reserved. JUDICATURE.DUKE.EDU

also others who have interacted with the judge in a professional setting: litigants (including self-represented), jurors, court staff, peer judges, law enforcement, probation officers, social workers, and witnesses, all of whom have perspectives to contribute. Within this broad base, however, surveys should be distributed only to those who have recently and personally interacted with the judge. Survey questions should focus on specific, readily observable behaviors rather than general attributes and should ask respondents to recall specific instances of those behaviors. Questions should also be tailored to specific audiences, placing the greatest emphasis on topics within the respondents' personal knowledge (such as judicial competence for lawyers, administrative capacity for court staff, and communication and impartiality for litigants).

The promised confidentiality of survey responses has caused much heartburn in practice. On the one hand, too many attorneys and court staff believe that their responses are not truly confidential, raising fears of judicial retribution should their identity be revealed. On the other hand, judges note that not knowing who responded makes it difficult for them to contextualize comments or act on feedback. Some survey respondents also hide behind the cloak of anonymity to submit comments that are racist or sexist, or which amount to threats or personal attacks. To promote a healthier exchange of survey responses, JPE 2.0 recommends that the identity of survey respondents remain confidential but that 1) attorneys first answer non-identifying questions about their experience and type of practice to provide context for their responses, 2) surveys be structured to reduce the risk of bias (by, for example, including "structured free recall" exercises that focus a respondent's thoughts on the behaviors and qualities of a specific judge), and 3) there be a clear process for defining and removing inappropriate comments before results go to the judge or commission, and that respondents be made aware that such comments will be removed.

Judges want more actionable feedback and more robust resources to support JPE implementation.

As noted above, a perennial challenge with surveys is low response rates, which can reduce confidence in the data. JPE programs should prioritize strategies to improve response rates, using technology (text messages, QR codes, etc.) to increase outreach and produce tailored messages to increase respondent buy-in.

Self-represented litigants, an increasingly large and important segment of court users, will require a multimodal approach to outreach and distribution, and JPE programs should pay attention to emerging platforms and strategies for capturing their perspectives.

Survey data should always be supplemented with additional information on a judge's performance, including case-management data, courtroom observation, decision and opinion review, self-evaluations, public input, and interviews with evaluated judges. This list is not necessarily exhaustive, and each JPE program will tailor data collection to its specific needs. To reduce error, judges should also be allowed to review their assessments and challenge aspects that they deem inaccurate or unfair.

Build a culture of and embrace judicial performance improvement as a key aspect of JPE. JPE often serves a dual role: informing the public about a judge's performance and promoting a judge's professional development. Though these goals sometimes can be in tension with each other, both are crucial components of an effective JPE program, and performance improvement should not be ignored. The JPE Perspectives Survey and JPE 2.0 convenings revealed that judges value feedback and consider the JPE process to be beneficial to their professional growth. For many judges, JPE evaluations are the only source of systematic feedback about their performance. At the same time, judges want more actionable feedback and more robust resources to support its implementation.

To this end, states should clearly communicate and demonstrate the importance of JPE to the professional development of the judiciary. Judges, evaluators, and the public should share the expectation that an evaluation is both a snapshot of the judge's performance at a particular point in time and a starting point for further professional growth. JPE programs should collaborate with judiciaries and judicial education offices to promote opportunities for judges to improve their performance following each evaluation. To foster this improvement, JPE and judicial education programs might collaborate to provide judges with personalized mentoring and wellness support.

Display institutional commitment to JPE. The convenings emphasized the need for strong institutional support to enable JPE programs to achieve objectives. States must demonstrate commitment to JPE through adequate funding and staffing and to fostering a culture of accountability and improvement. ▸

Published by the Bolch Judicial Institute at Duke Law. Reprinted with permission.
© 2026 Duke University School of Law. All rights reserved. JUDICATURE.DUKE.EDU

Institutional commitment starts with a stable and enduring program, ideally one legally grounded in legislation or the state constitution. It also requires stable and adequate funding to effectuate the goals of JPE, including the goal of professional development.

In addition to tangible resources, successful JPE programs require both symbolic and substantive commitment from the judiciary, legislators, and the bar. Judges should encourage those who appear before them to provide feedback. Legislators should raise awareness of JPE programs among their constituents and ensure that such programs are well-funded. The bar should remind attorneys to participate in evaluating judges.

Good faith support for JPE from each stakeholder group is needed to foster a culture of continuous judicial improvement.

Ensure and improve transparency for intended audiences. Transparency is a key component of public trust, but there is a delicate balance between transparency and performance improvement. Some approaches to JPE that emphasize transparency and public accountability (such as publishing details about a judge's evaluation) might disincentivize performance improvement by making judges feel defensive and thus less likely to use the feedback.

Other approaches, such as restricting the dissemination of evaluation results to those within the judiciary, create a safer environment for reflection but provide less transparency and public accountability. Each state will need to strike its own balance, taking into account the purpose and structure of its JPE program.

Regardless of where it lands on public dissemination of evaluations, each program should be clear with the public about how JPE works and what the evaluations mean. Common misunderstandings about the scope and rigor of the evaluations (especially the mistaken belief that JPE is a "rubber stamp" for the judiciary because most evaluations result in a retention recommendation) should be addressed with clear explanations and, at minimum, be accompanied by aggregate statistics reflecting the number of judges who have or have not met performance standards. In the same vein, JPE programs should ensure that judges have access to accurate information about the JPE process, including reasons for the commission's findings.

Further details on each of these recommendations can be found in the IAALS report.

NEXT STEPS

IAALS is working directly with states to help them implement JPE programs based on the JPE 2.0 model. Because JPE should be tailored to the needs of the judiciary and the public in each jurisdiction, the specific tools and approaches may differ slightly. States with current JPE programs are encouraged to assess their existing approaches, engage local stakeholders, and take action in line with these new recommendations. States that do not yet have JPE programs are encouraged to use the JPE 2.0 principles to craft a new program. In either case, the JPE 2.0 project has reiterated that open communication and transparency, stakeholder engagement, and focusing on the professional development of the judiciary are critical to enhancing public confidence in the legal system.

JPE is not an isolated exercise. The challenges facing JPE programs speak to broader challenges related to public trust in our democratic institutions. The fundamental principles underlying JPE — accountability, transparency, judicial independence, public civic engagement — are more vital than ever. The lessons learned from this work have applicability far beyond formal JPE programs, and these recommendations offer a roadmap for a more transparent and fairer judiciary.

IAALS is committed to ensuring that each state has access to the necessary tools and knowledge to build a judicial system that earns and maintains the public's trust.

Read more about JPE 2.0, IAALS recommendations, and next steps for states at ***https://iaals.du.edu/projects/judicial-performance-evaluation.***

JORDAN M. SINGER is a professor of law at New England Law Boston and chair of the IAALS JPE 2.0 Task Force.

DANIELLE KALIL is the director of Civil Justice and the Judiciary at IAALS.

Published by the Bolch Judicial Institute at Duke Law. Reprinted with permission.
© 2026 Duke University School of Law. All rights reserved. JUDICATURE.DUKE.EDU

1 *See Overview, Projects: Judicial Performance Evaluation 2.0*, Inst. for the Advancement of the Am. Legal Sys., https://duke.is/IAALS-JPE (last visited May 15, 2026) (noting JPE programs "focus on . . . holding judges accountable to high quality job performance while maintaining judicial independence, by focusing on politically neutral qualities like impartiality, case administration, and clarity of communication.").

2 *See* Malia Reddick & Rebecca Love Kourlis, *Judicial Performance Evaluation in* The Improvement of the Administration of Justice 155, 159 (Peter Koelling ed., 2016), https://duke.is/judicial-performanceevaluation ("JPE by state judicial branch entities took off in the 1980s, with six states — Connecticut, Utah, New Hampshire, Illinois, Colorado, and New Jersey — establishing programs. . . . The trend continued in the 1990s, with Arizona, Hawaii, Rhode Island, New Mexico, Florida, and Idaho adopting JPE. Massachusetts was close on these states' heels, enacting its program in 2001. [By] 2015, 17 states and the District of Columbia offer[ed] official programs for evaluating judicial performance.").

3 *See* Jordan M. Singer, Judicial Performance Evaluation in the States: The IAALS JPE 2.0 Pre-Convening White Paper 5 (2022), https://duke.is/jpewhitepaper (discussing defunded and discontinued JPE programs).

4 *See* Inst. for the Advancement of the Am. Legal Sys., Shared Expectations: Judicial Accountability in Context 19 (2006), https://duke.is/sharedexpectations ("The benefits of judicial performance evaluation were initially recognized in the 1970s. The first state-run JPE program began in Alaska in 1975.").

5 *See* Samuel W. Witwer Jr., *The Illinois Constitution and the Courts*, 15 U. Chi. L. Rev. 53, 58 (1947) (describing the Chicago Bar Association's "Bar Primary" process).

6 *See Errol E. Meidinger, Bar Polls: What They Measure, What They Miss*, 60 Judicature 468, 471 (1977) (noting that among a 1976 American Judicature Society comprehensive survey "[o]nly one poll . . . actually excluded lawyers with insufficient exposure to the candidates. This was done by checking court records and polling only those lawyers who had appeared before judges they were appraising.").

7 *See* Steven Flanders, *Evaluating Judges: How Should the Bar Do It?*, 61 Judicature 304, 304 (1978) (discussing the growth of the bar association poll industry).

8 *See* Richard L. Aynes, *Evaluation of Judicial Performance: A Tool for Self-Improvement*, 8 Pepp. L. Rev. 255, 264 –267 (1981) (describing the development of the early ABA programs for the evaluation of judicial performance).

9 *See generally* Am. Bar Ass'n, Black Letter Guidelines for the Evaluation of Judicial Performance (2005).

10 Reddick & Kourlis, *supra* note 2 and accompanying text.

11 *See*, *e.g.*, Rebecca Love Kourlis & Jordan M. Singer, *A Performance Evaluation Program for the Federal Judiciary*, 86 Denv. U. L. Rev. 7, 18–19 (2008) (discussing a pilot program in the Central District of Illinois).

12 *See* Reddick & Kourlis, *supra* note 2, at 156 ("Many states have 'official' programs for evaluating judicial performance that are authorized by constitution, statute, or court rule, and designed and implemented by state entities. 'Unofficial' JPE programs are also in place around the country, sponsored by bar associations, civic organizations, the media, or special interest groups.").

13 See Am. Bar Ass'n, *supra* note 9 ("All court systems should develop and implement a formal program for the evaluation of judicial performance.").

14 *See* Singer, *supra* note 3, at 3 ("By 2006 . . . nineteen states plus Puerto Rico and District of Columbia had instituted state-sanctioned judicial performance programs.").

15 *See id.* at 4–5 (noting that North Carolina and Idaho began JPE programs in the 2010s and discussing the defunding and deterioration of programs in Kansas, Tennessee, North Carolina, and Florida).

16 *See generally* Jordan M. Singer, *Attorney Surveys of Judicial Performance: Impressionistic, Imperfect, Indispensable*, 98 Judicature 20 (2014) (suggesting that surveys are essential tools for conducting judicial performance evaluations).

17 *See* Rebecca D. Gill, *Implicit Bias in Judicial Performance Evaluations: We Must Do Better Than This*, 35 Just. Sys. J. 301, 301 (2014) (showing "empirical evidence that gender and race bias still exist in attorney surveys conducted in accordance with the ABA's Guidelines"). See also Natalie Knowlton & Malia Reddick, Leveling the Playing Field: Gender, Ethnicity, and Judicial Performance Evaluation (2012), https://duke.is/level-the-playing-field.

18 *See* Singer, *supra* note 3, at 19 (noting "several judges expressed concern about low response rates for attorney and other surveys, especially because judges believed that attorneys with poor impressions of the judge would be more motivated to complete the survey and the overall data would be negatively skewed").

19 *See id.* at 24 ("Attorneys and some members of the public have complained about lack of transparency in JPE programs, especially in states in which they are asked to complete surveys but never learn the results of judicial evaluations.").

20 *See* Inst. for the Advancement of the Am. Legal Sys., Judges Aren't Sexy: Engaging and Educating Voters in a Crowded World 1 (2016), https://duke.is/engagingandeducatingvoters ("Voters can't find evaluation information or don't understand it when they do find it.").

21 *See* Singer, *supra* note 3, at 5 (discussing the discontinued and defunded JPE programs).

22 *See* Danielle Kalil, Jpe 2.0: Recommendations For Modernizing Judicial Performance Evaluation 2 (2025), https://duke.is/modernizingjpes (noting "it is important to modernize the JPE process to ensure it reflects the current needs and expectations of court users, judges, and the public").

23 *See* Danielle Kalil, *The Future of Judicial Performance Evaluation: Reinvigorating a Cornerstone of Public Trust*, Inst. for the Advancement of the Am. Legal Sys. (Nov. 30, 2023), https://duke.is/futureofjpe (discussing the virtual convenings "featuring a diverse group of participants . . . including judges, JPE commissioners, attorneys, court staff, and researchers").

24 *Id.*

25 For two thoughtful but limited efforts to collect judicial viewpoints on JPE programs, *see* Kevin M. Esterling & Kathleen M. Sampson, Judicial Retention Evaluation Programs in Four States: A Report with Recommendations 41–54 (1998); Inst. for the Advancement of the Am. Legal Sys., The Bench Speaks on Judicial Performance Evaluation: A Survey of Colorado Judges i–ii (2008).

26 *See* Danielle Kalil et al., Jpe 2.0: National Perspectives On Judicial Performance Evaluation 4 (2024), https://duke.is/perspectives (providing "Participating State Survey Administration Dates and Response Rates").

27 *Id.* at 6. Some states refer to judicial performance evaluation as judicial performance review, or JPR. The two terms were used interchangeably in the survey. We use the more common term "JPE" throughout this article.

28 *Id.*

29 *See id.* (noting that 63.6% of respondents agreed that the "JPE/JPR program increase[d] [their] judicial independence").

30 *Id.* at 5, 23.

31 *See id.* at 8 (summarizing judges' experiences with "[f]ear and [s]tress" related to JPEs).

32 Kalil, *supra* note 23; *see generally* Singer, *supra* note 3, at 5–8 (discussing evaluation criteria).

33 *See* Aynes, *supra* note 8, at 285 (noting substantial agreement upon judicial performance criteria in bar association surveys).

34 *See generally* Victor E. Flango & Thomas M. Clarke, Reimagining Courts: A Design for the Twenty-First Century (2014) (summarizing the changing roles and responsibilities of state courts and state judges in recent years).

35 Kalil, *supra* note 23; *see generally* Singer, *supra* note 3 (discussing self-represented litigants, technology changes, and case-management data).

36 *See generally* Pew Rsch. Center, Public Trust in Government: 1958-2025, https://duke.is/publictrust (2025) (analyzing public trust in government between 1958–2025); Logan Cornett & Natalie Anne Knowlton, *Public Perspectives on Trust and Confidence in the Courts* (2020), https://duke.is/publicperspectives (analyzing public perspectives on the courts through a qualitative lens).

37 Kalil, *supra* note 23.

38 *See generally* Kalil et al., *supra* note 26 (providing recommendations for modernizing JPE programs).