An Updated Blueprint for Judicial Performance Evaluation
Quality Judges is an initiative of IAALS dedicated to advancing empirically informed models for choosing, evaluating, and retaining judges that preserve impartiality and accountability. Through comprehensive analysis of existing practices and the collaborative development of recommended models, the Quality Judges Initiative empowers, encourages, and enables continuous improvement in processes for choosing, evaluating, and retaining judges.
# Table of Contents

Introduction .................................................... 1
The Commission .................................................. 2
  Authorization .................................................. 2
  Placement ....................................................... 2
  Scope .......................................................... 2
  Openness/Transparency ....................................... 3
  Commission Composition, Selection, and Terms ............... 4
  Code of Conduct for Commission Members ..................... 6
  Training ....................................................... 6
The Evaluation Process .......................................... 7
  Criteria ....................................................... 7
  Surveys ....................................................... 7
    Including Self-Represented Litigants ....................... 7
    Addressing Concerns about Low Response Rates .......... 8
    Limiting the Influence of Implicit Biases ............... 8
    Ensuring Anonymity of Respondents ..................... 11
Other Components ............................................. 11
  Case Management Data ..................................... 11
  Courtroom Observation .................................... 12
  Decision/Opinion Review .................................. 12
  Public Input .............................................. 13
  Self-Evaluation .......................................... 13
  Interview ................................................. 14
  Evaluating Appellate Judges ................................ 15
    “Appeal” by the Evaluated Judge ........................ 15
Recommendations for Reselection ................................. 16
  Dissemination of Results .................................. 18
  Funding ..................................................... 18
  Institutional Commitment to the Judicial Performance Evaluation Process .......... 19
Appendices ..................................................... 20
Judicial performance evaluation, or JPE, is the jewel in the crown of the O’Connor Judicial Selection Plan; when it works well, it balances the need for judicial impartiality and independence from external pressures with the opportunity to hold judges accountable for their on-the-job performance. Robust evaluation serves both to educate judges about ways in which their performance may be wanting and to provide meaningful information that voters, governors, and legislators can use in deciding whether to retain or reappoint judges.

But, for it to work, the process must have legitimacy. The judiciary must trust it, and the public must trust it. It cannot be a mere doffing of the hat to accountability: it must be thorough, objective, and credible.

Judicial performance evaluation has been around, at least in a few states, since the late 1970s. Currently, seventeen states and the District of Columbia have a judicial performance evaluation program authorized by constitution, statute, or court rule. In the last few years, one state (Kansas) has de-funded its judicial performance evaluation program and another (Tennessee) has discontinued such a program; but Virginia’s legislature reinstated funding for its program in 2015, and several states have made great strides in expanding and enhancing their programs in recent years.

A decade ago, one of the very first publications IAALS released was Transparent Courthouse: A Blueprint for Judicial Performance Evaluation. This well-received report provided principles, guidelines, and practical tools for establishing or enhancing a judicial performance evaluation program. Since the publication of Transparent Courthouse, we have continued to work to improve and expand judicial performance evaluation programs nationwide: conducting extensive research; holding conferences; conferring with judges, evaluators, and voters; drafting and redrafting surveys; and facilitating invaluable conversations and shared resources among administrators of judicial performance evaluation programs around the country through our JPE Working Group.

Based on our extensive work in this area over the last decade, the research we and others have conducted, and our interactions with our JPE Working Group, we have learned much about what works well—and what does not work as well—in evaluating judges’ performance, and we felt it was time to update the blueprint we offered in 2006. Transparent Courthouse Revisited: An Updated Blueprint for Judicial Performance Evaluation provides that update, offering a menu of recommended practices and tools for designing and implementing a judicial performance evaluation program that fosters legitimacy in the eyes of the public and the judges.

We offer a few caveats at the outset. First, the eighteen judicial performance evaluation programs in place around the country pursue a range of goals—from informing voters in judicial retention elections, to better equipping governors or legislators who reappoint judges, to promoting public trust and confidence in the judiciary, to simply encouraging judicial self-improvement. Many of the recommended practices we discuss here are directed particularly to those states with the weighty responsibility of informing judicial voters, but several of our recommendations are relevant to any program that endeavors to evaluate judicial performance.

Our second caveat is this: no single program includes all the elements we discuss here, and as such, the elements discussed here are aspirational and encouraged—and are not offered as potential criticisms of existing programs. Our intent is to identify and compile quality practices from programs currently in operation and offer them up as alternatives for consideration. Ultimately, whether a program is well accepted in the state in which it operates is influenced by the unique culture of that state. Our primary point is that there is always room for improvement—and always room to learn from the experiences of others.

Finally, the “blueprint” we offer here is based on the considered experiences of IAALS and several states working in judicial performance evaluation for many years now. Though well informed, these recommendations are not based on a systematic, quantitative assessment of the effectiveness of various elements of existing judicial performance evaluation programs, and we do not make such a claim.

The Commission

Particularly in states where the judicial performance evaluation program is charged with making recommendations to voters and others who reselect judges, a broad-based judicial performance evaluation commission should be established to implement and oversee the evaluation process. We offer here some guidance for the structure, selection and composition, and operation of such commissions.

Authorization

Most judicial performance evaluations are authorized by statute, though one program (Arizona) is constitutional and a handful of other programs are established in court rules. Statutory authorization provides a balanced approach, ensuring some degree of permanence coupled with the flexibility to adopt improvements as needed.

Placement

Several states have opted to place their judicial performance evaluation commissions within the judicial branch. In the interest of preserving judicial independence and limiting potential politicization of the evaluation process, this setup is preferable to housing the commission in the executive or legislative branch. However, to avoid even the appearance of judicial influence, the better option would be to create an independent office to administer the judicial performance evaluation program. Utah’s judicial performance evaluation commission, for example, is an independent governmental agency, housed purely for administrative purposes in the executive branch.

Where the commission is placed within the judicial branch, however, steps should be taken to ensure that the commission and its staff are structurally independent of the rest of the judiciary and the state court administrator. This may be accomplished by providing the commission with a separate line-item budget within the judicial branch budget, its own staff, and structural and decisional autonomy. Actual geographic distance from the offices of the court system may be helpful, as well.

Colorado’s authorizing legislation strives to ensure the independence from the judicial department and the legislature of the office of judicial performance evaluation and its executive director:

13-5.5-101.5. Office of judicial performance evaluation.

There is hereby established in the judicial department the office of judicial performance evaluation, referred to in this article as the “office”. The state commission on judicial performance established pursuant to section 13-5.5-102 shall oversee the office.

The state commission shall appoint an executive director of the office who shall serve at the pleasure of the state commission. The compensation of the executive director shall be the same as the general assembly establishes for a judge of the district court. The compensation paid to the executive director shall not be reduced during the time that a person serves as executive director. The executive director shall hire additional staff for the office as necessary and as approved by the state commission.

Scope

Currently, Arizona, Colorado, New Mexico, and Utah conduct both interim and retention-year evaluations of judges. Interim evaluations are consistent with one of the primary purposes of judicial performance evaluation programs: equipping judges on a more regular basis with the feedback they need for self-improvement. At the same time, interim evaluations can tax scarce resources, including the commission’s budget. Commissions should carefully consider whether conducting interim evaluations necessitates scaling back important aspects of retention-year evaluations. Special priority should be given to the resources available to publicize evaluation results in retention election years.
Whether or not interim evaluations are conducted throughout judges’ terms, new judges should have the benefit of an interim evaluation before they are first subject to retention or reappointment. Virginia’s judicial performance evaluation program, for example, requires such an interim evaluation:

§ 17.1-100. Judicial performance evaluation program.

B. … For any judge or justice elected or reelected on or after January 1, 2014, an interim evaluation of each individual justice or judge shall be completed during his term. Such interim evaluation shall be commenced by the judicial performance evaluation program no later than the midpoint of his term.

Openness/Transparency

Some commissions do not conduct public meetings, but opening the proceedings is likely to serve the goal of enhancing trustworthiness. Of course, even when commission meetings are open to the public, circumstances will arise in which executive session is warranted. In addressing when executive session is appropriate, Arizona’s rules of procedure balance well the public interest in the commission’s work with the need for open and frank discussion and the evaluated judge’s privacy interest:

Rule 2. Commission on Judicial Performance Review

e) Executive Session. The Commission shall meet in executive session with respect to any agenda item which would involve disclosure of matters made confidential by these rules, any other court rules, or by law. In addition, in order to promote open and frank discussion and accuracy in the performance evaluation process, the Commission shall meet in executive session at the time of: (1) discussion (not including voting) of the Commission’s finding as to whether a judge or justice “meets” or “does not meet” judicial performance standards; (2) presentation and discussion of a judge’s or justice’s written comment submitted in response to a finding that the judge or justice “does not meet” judicial performance standards; and (3) a judge’s or justice’s appearance before the Commission, provided, however, that an executive session in which a judge or justice appears shall be held prior to the public vote meeting. The Commission may meet in executive session at any other time upon a majority vote of the Commission members then in attendance. The substance of deliberations in executive session shall not be disclosed. All voting shall be in public session.

Currently, particularly in states where judicial performance is evaluated for self-improvement purposes only and/or states where judges are not subject to reselection, evaluation results are not always shared with the public. The programs in these states should consider the positive impact on public trust and confidence in the judiciary that providing at least aggregate results for each level of court may have. For example, New Hampshire conducts confidential performance evaluations of its judges for professional improvement but requires public reporting of summary data:

IV(B)(5) Report of Overall Evaluation Results.

The supreme court shall report the results of the judicial performance evaluation program in the annual report described in paragraph I above. This report shall include a summary of the overall evaluation results and all actions taken to correct inadequacies and deficiencies.
**Commission Composition, Selection, and Terms**

Many of the same principles for the selection, composition, and service of judicial nominating commissions outlined in IAALS’ *Goals and Principles for Judicial Nominating Commissions*² are applicable to judicial performance evaluation commissions as well.

The size of the commission should be determined based on the size of the jurisdiction the commission covers and the number of judges to be evaluated. Colorado is the only state that currently uses a statewide commission for appellate judges and 22 local commissions for trial judges, though Missouri only moved away from a similar setup in 2016. Local commissions may be more familiar with the expectations for and actual work of the judges they evaluate and can reduce the workload of a statewide commission. But there are disadvantages as well, including the administrative challenges and additional cost associated with multiple commissions.

Performance evaluation commissions must be broadly representative of the communities they serve—politically, ideologically, and demographically. Racial/ethnic and gender diversity among commission members should be strongly encouraged, and geographic diversity should be required.

It is critical that commission members be appointed by multiple authorities, to further enhance public confidence in the commission’s independence. Colorado and Utah, where the governor, legislative leaders, and either the chief justice or the supreme court chooses commission members, are exemplary models. Concerns have been raised that when the chief justice or the supreme court appoints commission members, those members may be biased in evaluating those justices and/or the judiciary as a whole. However, on balance, supreme court justices are well positioned to appoint members, including retired judges, who understand the court system and how it should work. Allowing justices to make a limited number of appointments serves the interests of the commission and the evaluation process. However, a majority of commission members should be appointed by entities outside the judicial branch.

Public members should serve on the commission in equal numbers to attorney members, and they may comprise a majority as in Arizona, Colorado, New Mexico, and Utah. The process of applying to serve on the commission should be open to the public.

Because of their firsthand familiarity with judicial roles and responsibilities, retired judges can make an invaluable contribution to the commission’s work; however, because of the potential for significant conflicts of interest, sitting judges should not serve on performance evaluation commissions.

After the initial selection of commission members, members’ terms should be staggered. This offers the advantages of preventing complete turnover in the commission’s membership, providing new members with the benefit of existing members’ experience, and ensuring rotation among appointing authorities. Term lengths of commission members should be four to six years, to enable commissioners to gain the valuable experience and expertise that will enhance the evaluation process and assist in the acclimatization of new members. Commission members should serve no more than two or three consecutive terms, depending on term lengths.

---

Utah’s commission exemplifies most of these principles, including ensuring political balance on the commission:


(1) There is created an independent commission called the Judicial Performance Evaluation Commission consisting of 13 members, as follows:

(a) two members appointed by the president of the Senate, only one of whom may be a member of the Utah State Bar;
(b) two members appointed by the speaker of the House of Representatives, only one of whom may be a member of the Utah State Bar;
(c) four members appointed by the members of the Supreme Court, at least one of whom, but not more than two of whom, may be a member of the Utah State Bar;
(d) four members appointed by the governor, at least one of whom, but not more than two of whom, may be a member of the Utah State Bar; and
(e) the executive director of the Commission on Criminal and Juvenile Justice.

(2) (a) The president of the Senate and the speaker of the House of Representatives shall confer when appointing members under Subsections (1)(a) and (b) to ensure that there is at least one member from among their four appointees who is a member of the Utah State Bar.

(b) Each of the appointing authorities may appoint no more than half of the appointing authority’s members from the same political party.

(c) A sitting legislator or a sitting judge may not serve as a commission member.

(3) (a) A member appointed under Subsection (1) shall be appointed for a four-year term.

(b) A member may serve no more than three consecutive terms.

(4) At the time of appointment, the terms of commission members shall be staggered so that approximately half of commission members’ terms expire every two years.

(5) When a vacancy occurs in the membership for any reason, the replacement shall be appointed for the unexpired term.
**Code of Conduct for Commission Members**

Commission rules of procedure should address the ethical responsibilities of commission members, including such areas as potential conflicts of interest, the appropriate extent of commission members’ political activity, their ex parte communications, and the confidentiality of the commission’s work. IAALS’ *Model Code of Conduct for Judicial Nominating Commissions* sets out several ethics provisions that are also relevant to members of judicial performance evaluation commissions. For example, provisions such as these may be applied equally well to judicial performance evaluation commissioners:

**Rule 1.1**

Members of judicial performance evaluation commissions hold positions of public trust. Commission members shall conduct themselves in a manner that reflects positively upon the judicial performance evaluation process and shall avoid partisanship or partiality in the consideration of applicants. Commission members shall not be influenced other than by facts or opinion relevant to judicial performance.

**Rule 2.1**

At the commencement of any commission meeting where the performance of judges will be considered, commission members shall disclose to the commission any relationship with a judge being evaluated (business, personal, attorney-client) or any other possible cause for conflict of interest, bias, or prejudice. A commission member shall recuse himself or herself if s/he might be unable to consider a judge being evaluated impartially and objectively.

**Rule 2.3**

No commission member shall hold any elected public office for which s/he receives compensation during his or her period of service. No commission member shall hold an official position in a political party or political campaign. No commission member shall make a campaign contribution on behalf of a judicial candidate.

**Rule 4.5**

Commission members are permitted, and encouraged, to communicate with the public regarding the commission’s role of evaluating judges and informing the public, in accordance with these Rules.

**Training**

New commissioners should complete a training program that addresses commissioner responsibilities, reviews the commission’s rules of procedure, and familiarizes members with the evaluation process. An important topic for such a training program is implicit biases, especially the potential for such biases to affect judicial evaluations and commissioner decisions. Additional training should be conducted periodically and as procedural rules are amended.

Colorado’s statute includes training members of the state and district commissions among the responsibilities of the executive director of the office of judicial performance evaluation, and its rules of procedure require biannual training:

**Rule 9. Training.**

The Office of Judicial Performance Evaluation shall provide training each retention year that is reasonably accessible and convenient to all commissioners. Each commissioner shall attend one training session, or an appropriate alternative as developed by the Office of Judicial Performance Evaluation, each year in which the commissioner is to evaluate a judge or justice.

---

The Evaluation Process

Criteria
States generally evaluate judges on criteria consistent with the recommendations in the American Bar Association's Guidelines for the Evaluation of Judicial Performance—i.e., legal ability, integrity and impartiality, communications skills, professionalism and temperament, and administrative capacity. Some programs also assess diligence and service to the legal profession and the public. Most surveys currently in use ask respondents about a range of aspects of each criterion, rather than simply asking respondents to assess judges on each overall criterion.

Judicial performance evaluation has a normative function. Judges will comport their behavior to what is being measured, just as any employee will internalize the performance criteria applicable to his or her job and seek to perform well on those criteria.

The “blueprint” proposed here takes no position on whether the ABA criteria, and how they are typically assessed, continue to be relevant and appropriate and whether they accurately and comprehensively capture the qualities that make a good judge. In the coming months, IAALS plans to tackle this question head on, revisiting the criteria used across the country and the assessment of performance against those criteria.

Surveys
The key component of any evaluation process is the survey that is disseminated to those who have interacted with judges in the courtroom and the office. In the past, surveys were in hard copy only, disseminated and returned by mail, with the attendant postage expenses. Today’s technology offers an alternative to that cumbersome and costly process. In order to keep costs down and to maximize response rates, survey respondents should be contacted via email regarding the opportunity to complete surveys, and surveys should be conducted online, with appropriate security controls. To facilitate this, commissions should work with court administration to acquire email addresses for as many court users, including self-represented litigants, as possible.

Including Self-Represented Litigants
Most states do a commendable job of surveying a broad swath of court users and others, including attorneys, other judges, court staff, jurors, litigants, witnesses, and state employees who regularly appear in court, such as peace and probation officers and social workers. However, there is a growing body of court users whose views are not sufficiently sought or addressed: self-represented litigants. Self-represented litigants are not only entitled to have their views heard, but they also have distinct needs as they navigate the court process and interact with judges. It is imperative that judicial performance evaluation programs, and courts themselves, begin to grapple with this reality. Few states currently take the necessary steps to ensure that self-represented litigants have the opportunity to weigh in on their experiences, including the performance of the judges before whom they appear. To ensure that this occurs, specialized questionnaires and procedures for identifying and surveying respondents may need to be instituted.

We acknowledge that this is not an easy task, and that some state programs have attempted, and been frustrated by, efforts to incorporate the perspectives of self-represented litigants. At the same time, most of us would agree that the effort is an essential one. More and more individuals have email access, and inviting self-represented litigants to complete an online survey is something that would enhance both the effort to obtain broad-based feedback on judges’ performance and the legitimacy of the process.

The Access and Fairness Survey that is included among the *Trial Court Performance Measures* developed by the National Center for State Courts provides one model for a survey of self-represented litigants and is reproduced as Appendix A. However, as the survey includes questions about both court resources and individual judge performance, this survey may not be well suited for a judicial performance evaluation program.

**Addressing Concerns about Low Response Rates**

One of the most serious criticisms of judicial performance evaluation programs is the low survey response rate and the resulting lack of generalizability of results. Response rates are a problem for nearly all survey research, but even in the absence of high response rates, survey data can provide valuable insight. Also, response rates are less important than the representativeness of the sample. It is entirely possible to have a low response rate and a representative sample; conversely, it is also possible to have a high response rate and a non-representative sample.

One option for addressing concerns in this area is to include questions about the nature and extent of the respondents’ interaction with the judge in the survey. This information can then be summarized for each judge’s respondents and included in the evaluation report. This will give both evaluated judges and the commission a sense of whether those who responded are representative of all potential respondents, regardless of the response rate. Alaska, for example, asks about attorney respondents’ basis for evaluating the judge (including in the evaluation report only the responses of those who had direct professional contact), type of legal practice, years of practice in the state, gender, types of cases handled, and location of practice. Similar questions are asked of other respondents.

Most judicial performance evaluation programs do not claim that the results are based on a random sample of respondents whose views are representative of—or generalizable to—the population of potential respondents with a specified margin of error or confidence level. Rather, they state that the results are based on responses received from members of the pool of potential respondents who took the time to complete the survey. Judges should consider reminding those with whom they interact that they will likely be receiving surveys, that they are encouraged to fill out the survey, and that their feedback is welcomed.

It is worth noting here that one of the reasons a judicial performance evaluation commission is essential is that survey results alone may not be a sufficient window into a judge’s performance on the bench. As discussed later in this report, the commission should collect and evaluate other information and exercise their own independent judgment to reach conclusions that will help both the judge and those who reselect judges.

**Limiting the Influence of Implicit Biases**

Survey respondents may have implicit biases that creep into their evaluations. The judicial performance evaluation process should be structured to minimize the possibility that implicit biases based on gender, race, and/or ethnicity will influence assessments of judges. One way to explore whether implicit biases are a factor is to examine whether survey respondents systematically score women and/or minority judges lower than white male judges. IAALS did just that in 2012, with *Leveling the Playing Field: Gender, Ethnicity, and Judicial Performance Evaluation*, and found small differences in evaluation scores based on gender and race/ethnicity for survey questions related to legal ability, communication skills, and temperament.

Based on this research, as well as work that others have done, IAALS made recommendations for combating the potential impact of implicit biases in judicial evaluations. First, we recommended that judicial performance evaluation programs be broad-based, particularly in terms of the types of respondents who are surveyed. The studies that initially sounded the alarm regarding implicit bias and judicial performance evaluation were based on surveys of attorneys only. The IAALS study examined programs that also included non-attorneys, other judges, and jurors as respondents, and found the largest differences based on gender and race/ethnicity in the surveys of attorneys. (To that end, we called for better education of attorneys on the potential for implicit biases to affect their judgments.)

---


Second, IAALS recommended that evaluation programs take care to ensure that only respondents who have had recent professional contact with the evaluated judge are surveyed. A person who has had actual contact with a judge is more likely to assess his or her performance on the basis of that interaction, not on the basis of stereotypes. On the other hand, someone who does not have direct professional experience with the judge may be more susceptible to implicit biases. This is a particular concern in programs that are administered by state bar associations, where all attorneys in the state may receive evaluation surveys. To address this concern, the Missouri Bar’s program poses these screening questions at the outset of its attorney surveys:

**FOR APPELLATE JUDGES:**

Please check the following statements that apply to you (may check more than one):

- I use appellate opinions in my practice.
- I have appeared before this judge in oral argument.
- This judge has written an opinion in one of my cases.
- None of the above

In the most recent case in which you appeared before this judge, did you represent the winning party or the losing party?

- Winning party
- Losing party
- Mixed outcome
- Case not yet decided
- N/A

**FOR TRIAL JUDGES:**

Do you have direct and personal knowledge of the behavior of Judge X based on your cases before Judge X during the judge's current term of office?

- Yes
- No

Please estimate the number of your cases that Judge X has presided over during the judge's current term of office.

- 0-5
- 6-10
- More than 10

Please estimate the number of your cases before Judge X that went to trial and decision during the judge's current term of office.

- 0-5
- 6-10
- More than 10

Of your cases before Judge X that went to trial and decision during the judge's current term of office, the rulings were...

- Primarily in your party’s favor
- About equally in favor of and against your party
- Primarily against your party
- N/A
Finally, IAALS recommended strongly that evaluation surveys be developed in consultation with experts in the field of job performance evaluation and survey design. In 2010, working with the Illinois Supreme Court Judicial Performance Evaluation Committee, the National Center for State Courts underwent a rigorous process of expert consultation, pilot testing, and cognitive interviews to craft attorney and court staff surveys for evaluating the performance of trial judges. IAALS replicated this process in 2013 in creating model surveys for attorneys, trial judges, and court staff to use in evaluating appellate judges.

IAALS learned some important lessons from these survey development projects. For example, survey experts recommend that a structured free-recall (SFR) component be incorporated in the survey. The SFR component calls on respondents to complete a memory-based task before beginning the actual survey—a proven technique in minimizing potential respondent biases. The SFR component asks respondents to recall positive and negative behaviors observed during their appearance before the evaluated judge, along stated performance dimensions, and to describe three of each. The structured-free recall component looks like this in IAALS’ model surveys for evaluating appellate judges:

**SECTION II: STRUCTURED FREE RECALL**

Research has shown that people make better and more accurate performance evaluations when they take a few minutes to think about specific aspects of the person's performance rather than simply relying on their general impressions of the person.

*To help you make a better performance evaluation, please take a few moments to recall some positive aspects of the evaluated judge's performance. In your experience with the judge, what did s/he do well? If it is helpful in organizing your thoughts, you may list these positive aspects here. Anything you record here will be discarded once you submit your survey; it will not be shared with the evaluated judge or the evaluation commission.*

1. 
2. 
3. 

*Now, please take a few moments to recall some negative aspects of the evaluated judge's performance. In your experience with the judge, what did s/he do poorly? If it is helpful in organizing your thoughts, you may list these negative aspects here. Anything you record here will be discarded once you submit your survey; it will not be shared with the evaluated judge or the evaluation commission.*

1. 
2. 
3. 

The Missouri and North Carolina bar association programs and the Virginia program adopted the SFR recommendation, in conjunction with screening questions for attorney respondents, and the Missouri and North Carolina programs saw a decline in concerns about implicit bias in the next election cycle. (The Virginia program is too new at this point to allow such an assessment.)

Survey experts also made recommendations regarding the content and structure of the survey itself, to limit the potential for implicit biases to come into play. For starters, question wording should reflect readily observable behaviors rather than general attributes or performance criteria. For instance, survey questions such as “The judge’s ruling cited applicable substantive law” and “The judge writes opinions that clearly set forth any rules of law to be used in future cases” are preferable to survey items such as “Legal knowledge” or “The judge is competent in the law.”

The response scale is important as well. When respondents are asked to rate aspects of a judge’s behavior as unacceptable, deficient, acceptable, good, or excellent, or to assign a grade of A, B, C, D, or F, each respondent must define for him or herself what the available responses mean, and different people are likely to have different standards regarding what
behaviors are unacceptable, or deserving of a C grade rather than a B. Using a frequency-based response scale, ranging from “Never” to “Every Time” (as in the National Center for State Courts surveys for Illinois judges) or a Likert scale of agreement/disagreement (as in the IAALS surveys for evaluating appellate judges) is preferable.

The surveys developed by the National Center (for evaluating trial judges) and IAALS (for evaluating appellate judges) are reproduced in Appendix B and Appendix C, respectively. (We offer two notes regarding the National Center surveys: first, the Illinois evaluation program for which the surveys were created is intended solely to promote judicial self-improvement, and as such, survey results are not made public; and second, judges subject to evaluation submit a limited number of names of attorneys and court personnel to complete the surveys, so the surveys are lengthier and more detailed than surveys in other programs where everyone who interacted professionally with a judge may be surveyed.)

Ensuring Anonymity of Respondents
To obtain candid feedback from those who evaluate judges, survey respondents must be confident that their responses are anonymous—that the feedback they provide will not, and indeed cannot, be attributed to them. This is another aspect of the survey process where consulting with experts is essential. Anonymity may be a particular concern where survey respondents take advantage of the opportunity to provide narrative comments about judges’ performance. Respondents must have assurance that they will not be identified as the source of such comments, placing a special responsibility in the commission to remove any potentially identifying features before deciding to share comments with the evaluated judge. Survey respondents do worry about retribution from judges for offering critical feedback and tell stories of such retribution allegedly occurring; such concerns, whether well founded or not, have the potential to undermine the evaluation process.

Other Components
The strongest judicial performance evaluation programs utilize a range and variety of information to achieve a fair, accurate, and broad-based assessment of judges’ on-the-job performance.

Case Management Data
Case management is an essential skillset for judges, and many states examine case processing data and compliance with case-under-advisement time standards as part of the evaluation process. It is important that this assessment focuses on the elements of case management over which the judge has authority, such as how quickly he or she rules on pending motions. Another relevant data point may be how often the judge continues cases or hearings sua sponte. Examining overall time-to-disposition of cases may or may not be reflective of the judge’s performance because there may be a variety of factors at work over which the judge has no control. Similarly, using that measurement too heavily risks incentivizing a “rocket docket” mentality and ultimately may encourage judges to push cases toward early resolution, whether or not that resolution is the right one for the case.
Utah’s program defines the case-under-advisement standard in the following manner:

**Rule 3-101. Judicial Performance Standards**

(1) Case under advisement standard. A case is considered to be under advisement when the entire case or any issue in the case has been submitted to the judge for final determination.

(2)(A) A justice of the Supreme Court demonstrates satisfactory performance by circulating not more than an average of three principal opinions per calendar year more than six months after submission with no more than half of the maximum exceptional cases in any one calendar year.

(2)(B) A judge of the Court of Appeals demonstrates satisfactory performance by:

(2)(B)(i) circulating not more than an average of three principal opinions per calendar year more than six months after submission with no more than half of the maximum exceptional cases in any one calendar year; and

(2)(B)(ii) achieving a final average time to circulation of a principal opinion of not more than 120 days after submission.

(2)(C) A trial court judge demonstrates satisfactory performance by holding:

(2)(C)(i) not more than an average of three cases per calendar year under advisement more than two months after submission with no more than half of the maximum exceptional cases in any one calendar year; and

(2)(C)(ii) no case under advisement more than six months after submission.

**Courtroom Observation**

Two states—Colorado and Utah—incorporate courtroom observation in their evaluation program, either formally or informally. Utah is a state with a formal, and quite sophisticated, process that gives members of the public a chance to participate in the evaluation of judges. The judicial performance evaluation commission first trains volunteers on the value of procedural fairness in the court context, and volunteers then observe judges in court. Observers record both what they see and their personal responses to what they see as it relates to procedural fairness. The courtroom observation reports are shared with judges and are included in the judge’s evaluation report.

Even where courtroom observation takes place informally, some systemization may be imposed to ensure that all judges are observed on separate occasions by at least two observers (usually commission members) during the evaluation period and that observers are assessing judges on the same qualities. The latter may be accomplished through a checklist or questionnaire, with space for narrative comments, which can then be shared with commission members.

**Decision/Opinion Review**

Another important component of a performance evaluation program is the review of opinions for appellate judges, and written decisions for trial judges. Decisions and opinions should be evaluated based on legal reasoning and analysis, fairness, and clarity, without regard to the particular outcomes reached. As with courtroom observation, for the sake of consistency, fairness, and quality, commission members should use some sort of template in reviewing judges’ opinions that can then be shared with fellow commissioners. Colorado’s rules of procedure offer these guidelines for reviewing decisions and opinions:
Rule 10. Trial Judge Evaluations

(e) Each trial judge being evaluated shall submit to the district commission not less than three decisions he or she issued, including, if applicable, one of which was reversed on appeal, together with the reversing opinion, if applicable. The judge may choose written or transcribed decisions for submission. Each district commission shall review the three decisions or transcripts and any others authored by the trial judge that the commission in its discretion may select for compliance with the statutory criteria for legal knowledge, thoroughness of findings, clarity of expression, logical reasoning, and application of the law to the facts presented. All decisions and opinions submitted or reviewed shall have been issued during the judge's current term.

Rule 11. Appellate Judge Evaluations

(e) Each appellate judge or justice shall submit to the state commission five opinions he or she authored, including both civil and criminal cases. These opinions shall include, if applicable, at least one separate concurrence or dissent, at least one unpublished opinion, and at least one opinion which was reversed on appeal, together with the reversing opinion. The state commission shall review the five opinions and any others authored by the appellate judge or justice that the commission in its discretion may select for compliance with the statutory criteria for legal knowledge, adherence to the record, clarity of expression, logical reasoning, and application of the law to the facts presented. All opinions submitted or reviewed shall have been issued during the appellate judge or justice’s current term.

The Missouri Bar program incorporates a similar review procedure, going so far as to identify judges whose decisions/opinions are being evaluated by number only and removing any identifying features to ensure that commission members’ assessments are objective and unbiased.

Public Input

Commissions should hold at least one public hearing per year, with multiple hearings around the state in an election year, to receive feedback on judicial performance and the evaluation process as a whole. Such hearings should be widely publicized well in advance of the dates on which the hearings will be held. Public hearings will provide another opportunity for litigants who represented themselves to weigh in on their court experiences. The Alaska Judicial Council describes its public hearings as follows on its website:

**Public Hearings**

The Council holds statewide public hearings for all judges standing for retention using the legislature’s teleconference network and public meeting rooms. Subject to available funding, the Council advertises these public hearings in statewide newspapers to encourage public participation. Public service announcements on radio and television stations encourage public participation. Public hearings give citizens a valuable opportunity to speak out about their experiences with judges. They also provide a forum in which citizens can hear the opinions of others. The Council tries to balance all the information it receives from all sources.

The public should also have the opportunity to submit anonymous feedback online regarding the performance of individual judges. Commission staff can “clean” such comments to remove any potentially identifying information before sharing with commission members and the evaluated judge about whom the comments were made. This practice does run the risk of eliciting biased or uninformed opinions, and such input should be given less weight than surveys and other objective data; however, if, for example, there are negative comments about a particular judge from a number of sources, those comments should at least prompt further inquiry by the commission.

Self-Evaluation

In many evaluation programs, judges complete a self-evaluation. Ideally, a portion of this self-evaluation will replicate some or all the questions posed in surveys of those who interact professionally with judges. This will allow judges the opportunity for a “reality check” as to whether their perceptions of their own performance comport with those of court users, colleagues, and staff. Judges may view the self-evaluation as perfunctory or self-serving, but when both the judge and the commission take it seriously, it can be beneficial to both.
Interview

Once the evaluation process is complete and a preliminary evaluation report has been prepared, the report should be shared with the evaluated judge. The commission should then hold an interview with each judge, at which time the judge can respond to the evaluation results, the commission can discuss areas of concern identified in the evaluation, and both parties can agree on ways to address any concerns. Even in states like Arizona that hold open meetings, the interview is conducted in executive session, and appropriately so.

Arizona’s program also takes the interview a step further, providing for a meeting with a three-member conference team that will assist each judge in developing a performance improvement plan:

**Rule 4. Conference Teams**

During each mid-term and retention election performance review period of a judge or justice, the Commission shall arrange for a conference between each judge or justice and a Conference Team. The purpose of this conference shall be to assist in identifying aspects of the judge’s or justice’s performance that may need improvement and to help the judge or justice to develop plans for self-improvement. The activities and operations of the Conference Teams shall be governed by the following provisions:

(a) Composition. Each Conference Team shall be appointed by the Chairperson of the Commission or his or her designee and shall be composed of a member of the public, an attorney who is a member of the State Bar of Arizona, and a judge or justice (active or retired). No more than one member of a Conference Team may be a member of the Commission.

(b) Organization. The Conference Team members shall organize themselves as meets their needs in order to conference with the judge(s) assigned to that team.

(c) Terms. A Conference Team may review more than one judge or justice during any review period. Conference Team members shall be recruited to serve for each judicial review cycle and service will terminate at the end of the specific review cycle.

(d) Meetings. Meetings shall be at the call of the Conference Team. All meetings shall be confidential. No meeting shall take place unless all three (3) members are present.

(e) Self-Evaluation Form. Prior to meeting with the Conference Team, each judge or justice shall complete a self-evaluation form approved by the Commission reflecting his or her perception of his or her performance as to each judicial performance criterion. The completed self-evaluation form is confidential and plays no role in the evaluation/retention process. It shall be furnished only to the Conference Team before its meeting with the judge or justice, and then to his or her Presiding Judge or Chief Judge, and to the Chief Justice, along with the self-improvement plan described in Paragraph (h) below.

(f) Peremptory Challenge. Each reviewed judge or justice shall have the right to peremptorily challenge one member of the Conference Team. The peremptory challenge shall be filed with the office of the Commission within 5 days of actual notice to the judge or justice of the members of the Conference Team. Where necessary, the Chairperson of the Commission shall rule upon any questions under this subparagraph.

(g) Conference Team Report. A written plan for self-improvement shall be developed at the conference and, after being put into final form, signed by the judge or justice and the Conference Team members. In connection with development of the self-improvement plan, the judge or justice and the Conference Team shall consider previous and current survey results and narrative comments, the previous self-improvement plan, and objective data which demonstrates completion of the previous plan. The self-improvement plan shall be distributed only to the judge or justice being reviewed, to his or her presiding judge or chief judge, and to the Chief Justice. In addition, the self-improvement plan, with the name of the judge or justice redacted, may be distributed to the Administrative Office of the Courts for use in development of judicial education programs. Neither the Conference Team Report nor the self-improvement plan shall be distributed to the Commission or used in the Commission’s deliberations as to whether a judge or justice “meets” or “does not meet” judicial performance standards.
EVALUATING APPELLATE JUDGES

In some states, programs for evaluating the performance of appellate judges have largely been patterned after programs for trial judges, though these judges’ roles and responsibilities differ significantly. Perhaps the most significant difference between appellate and trial judges is their work product; while trial judges hold open court proceedings and conferences and make rulings throughout the course of a trial, an appellate judge’s primary output is the written opinion. Programs must take these differences into account both in designing evaluation programs for each type of judge and in laying out the components of the evaluation process for each type of judge in the authorizing legislation or court rules.

With Recommended Tools for Evaluating Appellate Judges, IAALS provided surveys about appellate judges to be completed by attorneys who appear before them and use their opinions, trial judges who apply their rulings, and court staff who work closely with them; guidelines and templates for reviewing appellate judges’ written opinions for legal analysis and reasoning, clarity, and fairness; and a self-evaluation tool that allows judges to assess their own strengths, weaknesses, and overall performance.

“APPEAL” BY THE EVALUATED JUDGE

Performance evaluation programs must provide judges the opportunity to review the evaluation report and any overall assessment before this is made public and to challenge aspects that they deem inaccurate or unfair. This need not occur through a formal appeal process but may be accomplished simply through communications between the judge and the commission. If the judge’s concerns are not taken into account in revising the report and assessment, the judge should have the opportunity to make a short statement regarding his/her own perspective to be included in the publicized report.

Colorado’s rules of procedure describe a process that has worked well over the years:

Rule 13. Narratives

(b) The judge or justice being evaluated may respond in writing to the draft narrative within seven days of receipt of the draft. The judge or justice may provide feedback on or corrections to the draft narrative language, and may request an additional interview. Any additional interview shall be held within fourteen days of the request. The commission may revise the draft narrative, and shall provide the judge or justice with the final narrative within fourteen days following the written response or additional interview.

(c) Any commission issuing a “do not retain” or “no opinion” recommendation shall, at the judge or justice’s request, include a response from the judge or justice of not more than 100 words. The judge or justice shall have seven days from receipt of the commission’s final recommendation and narrative to submit the 100 word response to the chair of the commission or the executive director of the Office of Judicial Performance Evaluation, who will forward the response to the commission. The commission may then change its vote count or revise the narrative, and shall provide the judge or justice with the final narrative within seven days following the receipt of the response.

New Mexico’s program employs a similar procedure:

28-206. Evaluations and interviews.

E. Response to narrative profile. Any judge being evaluated may respond to a draft of a narrative profile, in writing, within ten (10) days of receipt of the draft. A response must be directed to the chair of the commission. If the responding judge requests an additional interview with the commission, the judge may be given an opportunity to meet with the commission to address the contents of the narrative profile. The commission may redraft the narrative profile prior to releasing it to the public.

Recommendations for Reselection

A fundamental responsibility of the judicial performance evaluation commission is to compile findings from the various evaluation tools and sources of information and issue an overall assessment. To limit the role of discretion, this assessment should be based largely on whether judges comply with clearly articulated performance standards, established by statute or court rule, and tied to the evaluation tools in use.

When performance standards are clear, the function of the judicial performance evaluation commission is clearer as well. Namely, the commission determines whether a particular judge meets or does not meet those standards.

Utah’s program, where performance standards are set out in both statute and administrative rule, again provides a good model:

**78A-12-205. Minimum performance standards.**

(1) The commission shall establish minimum performance standards requiring that:

(a) the judge have no more than one public reprimand issued by the Judicial Conduct Commission or the Utah Supreme Court during the judge's current term; and

(b) the judge receive a minimum score on the judicial performance survey as follows:

(i) an average score of no less than 65% on each survey category as provided in Subsection 78A-12-204(7); and (ii) if the commission includes a question on the survey that does not use the numerical scale, the commission shall establish the minimum performance standard for all questions that do not use the numerical scale to be substantially equivalent to the standard required under Subsection (1)(b)(i).

**R597-3-4. Minimum performance standards.**

(1) In addition to the minimum performance standards specified by statute or administrative rule, the judge shall:

(a) Demonstrate by a preponderance of the evidence, based on courtroom observations and relevant survey responses, that the judge’s conduct in court promotes procedural fairness for court participants.

(b) Meet all performance standards established by the Judicial Council, including but not limited to:

(i) annual judicial education hourly requirement;

(ii) case-under-advisement standard; and

(iii) physical and mental competence to hold office.

(2) No later than October 1st of the year preceding each general election year, the Judicial Council shall certify to the commission whether each judge standing for retention election in the next general election has satisfied its performance standards.
According to the website of the Arizona Commission on Judicial Performance Review, performance standards for Arizona judges are tied, at least in part, to survey results:

To determine whether a judge or justice “Meets” or “Does Not Meet” judicial performance standards, Commission members compare the compiled survey data of a judge or justice against two threshold standards. For each category of judicial performance, judges are rated and scored as “Superior” 4 points, “Very Good” 3 points, “Satisfactory” 2 points, “Poor” 1 point, and “Unacceptable” 0 points. First, if a judge or justice has an average score of 2.0 or less for any category from any group of respondents, the judge does not meet the threshold standard. Second, if 25 percent or more of any group of respondents rate the judge or justice as “unacceptable” or “poor” in any category, the threshold standard has not been met. The categories of “legal ability”, “integrity”, “communication skills”, “judicial temperament”, “administrative performance” and “administrative skills” are subject to the threshold standards. “Settlement activities” is not because it is very difficult to evaluate.

The original _Transparent Courthouse_ proposed the following benchmarks:

At minimum, an average performance on at least 80% of all survey questions (“average performance” meaning, for example, a score of 3.0 on a 1-5 scale, or at least 7% of respondents answering “yes” to a yes/no question);

For trial judges, no cases with issues under advisement more than 90 days, unless the judge's particular docket assignment justifies exceptions;

For appellate judges, the authorship of a proportionate number of opinions authored by that court on average in a given calendar year, taking into account both particularly complex cases and concurring or dissenting opinions authored during the same period;

All or nearly all written opinions clearly and accurately describe the relevant facts and applicable law, and clearly state the court’s order; and

No findings by a body charged with judicial discipline that the judge has violated the applicable code of judicial conduct.

While these recommended practices call in the strongest terms for commissions to offer an assessment to voters and others who reselect judges of each judge's fitness for reselection, they do not take a position on the form that assessment should take. Two options exist: issuing a “Retain” or “Do Not Retain” recommendation regarding each judge, as do commissions in Alaska, Colorado, New Mexico, and Utah, or making a determination of whether or not each judge “Meets” or “Does Not Meet” judicial performance standards, as do the Arizona and Missouri commissions.

A potential advantage to indicating whether judges meet performance standards is that it reinforces in the minds of voters that the evaluation process exists within the context of the standards applied in that state, and does not purport to address factors that are not included in the standards. In addition, commissions in Alaska and Colorado have faced allegations of improperly using state funds to engage in electioneering because they issue a retention recommendation. The “Meets” or “Does Not Meet” language may be less susceptible to being interpreted as the commission telling voters how to vote.

Commissions should also make public the vote count of commission members with respect to the number that agreed and disagreed with the commission's assessment of individual judges' fitness for reselection.
**Dissemination of Results**

One of the greatest challenges that judicial performance evaluation programs face is getting evaluation results in the hands of those responsible for judges’ reselection. Sharing results with governors or legislators who reappoint judges is fairly straightforward, but communicating with voters is far more difficult. IAALS recently took on this issue and, in collaboration with a political and communications consultant, offered a range of strategies, templates, and messaging for public dissemination, laid out in detail in *Judges Aren’t Sexy: Engaging and Educating Voters in a Crowded World.*\(^8\) The publication presents some inexpensive options, such as the purchase of Facebook and Google AdWords ads, strategic use of earned media, and succinct and understandable presentation of evaluation results, but it is still important that a meaningful portion of the program’s budget be dedicated to the voter communication task.

**Funding**

For judicial performance evaluation programs to achieve legitimacy, they must have ongoing, adequate, and secure funding. The use of technology in disseminating and conducting surveys can reduce some costs, but the process must still have adequate staffing and support. Also, it is essential that two aspects, in particular, of the commission’s work be prioritized in terms of funding. First, surveys of court users and others are the central piece of the evaluation process, and their credibility and validity must be unassailable. To accomplish this, the commission and its staff should consult with survey research experts on all aspects of survey administration, and funds should be dedicated to this purpose. Second, sufficient funding should be allocated to informing voters that evaluation programs are in place, that evaluation information is easily accessible and understandable, and that the evaluation process is credible and meaningful. Lower cost options for accomplishing this are available, but there is no alternative to strategic paid advertising and outreach; the program’s budget must enable this. No matter how effective an evaluation program is, the effort is largely for naught if voters are not aware of it and how to use it.

States currently employ a range of funding mechanisms for their judicial performance evaluation programs. The now-defunct Kansas program was funded by a portion of docket fees, though the legislature had to appropriate these moneys for the program. Other states’ programs are funded through a line-item in the judicial branch’s budget, and a few programs periodically obtain one-time government grants for specific purposes, such as technical assistance and public outreach.

New Mexico has a designated, non-reverting fund for its judicial performance evaluation program:

**34-9-18. Judicial Performance Evaluation Fund; Created.**

A. The “judicial performance evaluation fund” is created in the state treasury to be administered by the administrative office of the courts. The fund shall consist of appropriations, gifts, grants, donations and bequests made to the fund. Income from the fund shall be credited to the fund. Balances in the fund shall not revert to the general fund at the end of any fiscal year.

---

The legitimacy of a judicial performance evaluation program rests on the shoulders of those who strongly believe in it and actively support it. 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 judicial performance evaluation to the judicial branch, and assist judges in using evaluation results to improve their performance. Individual judges need to encourage those who appear before them to provide performance feedback, and be open to the feedback they receive. Legislators should raise awareness of judicial performance evaluation programs among their constituents, and ensure that such programs are adequately and securely funded. The bar must remind attorneys of their responsibility to participate in evaluating the judges before whom they have appeared, and assist in publicizing evaluation results. Members of the public, both laypersons and attorneys, should consider applying to serve on a judicial performance evaluation commission; they should also take the necessary steps to inform themselves as judicial voters and encourage friends and family to do the same.

The judicial performance evaluation commission’s commitment is to establish and administer an evaluation program that is transparent, fair, thorough, and robust. By considering the options set forth in this *Updated Blueprint for Judicial Performance Evaluation*, such a program is within reach.
APPENDICES

Available Online

APPENDIX A:

National Center for State Courts Access and Fairness Survey

Included in the Trial Court Performance Measures developed by the National Center for State Courts is an Access and Fairness Survey. The Access and Fairness Survey provides one model for a survey of self-represented litigants. (Note, however, that as the survey includes questions about both court resources and individual judge performance, it may not be well suited for a judicial performance evaluation program.)


APPENDIX B:

Supreme Court of Illinois Judicial Performance Evaluation Program

In 2010, working with the Illinois Supreme Court Judicial Performance Evaluation Committee, the National Center for State Courts underwent a rigorous process of expert consultation, pilot testing, and cognitive interviews to craft attorney and court staff surveys for evaluating the performance of trial judges. (Note that in the Illinois program, judges subject to evaluation submit a limited number of names of attorneys and court personnel to complete the surveys, so the surveys are lengthier and more detailed than surveys in other programs where everyone who interacted professionally with a judge may be surveyed.)


APPENDIX C:

IAALS Model Surveys for Appellate Judicial Performance Evaluation

In 2013, IAALS replicated the process used in Illinois by the National Center for State Courts in creating model surveys for attorneys, trial judges, and court staff to use in evaluating appellate judges. (As part of that project, IAALS also offered recommendations for reviewing the written opinions of appellate judges and opinion review templates for attorney and non-attorney reviewers.)
