

LEVELING THE PLAYING FIELD:

Gender, Ethnicity, and Judicial Performance Evaluation



“

BECAUSE IMPLICIT BIAS IS A SERIOUS
ISSUE THAT HAS THE POTENTIAL TO
IMPACT JPE—AND TO DO SO IN WAYS
THAT ARE DIFFICULT TO DETECT—
BOTH EXISTING AND DEVELOPING
JPE PROGRAMS SHOULD STRIVE TO
MINIMIZE THE POTENTIAL IMPACT OF
SUCH BIAS TO THE EXTENT POSSIBLE.

”

LEVELING THE PLAYING FIELD:

Gender, Ethnicity, and Judicial Performance Evaluation

NATALIE KNOWLTON

Manager, IAALS Quality Judges Initiative

&

MALIA REDDICK

Director, IAALS Quality Judges Initiative

ACKNOWLEDGEMENTS

IAALS would like to thank the American Judicature Society (AJS) and AJS Executive Director Seth Andersen, who worked closely with IAALS in the early stages of this study to define the parameters and goals of the analysis. AJS Research Assistants also participated in state data collection.

IAALS also thanks the participating state judicial performance evaluation coordinators and commissions for sharing raw data with us and lending transparency to their programs and processes, specifically: Larry Cohn of the Alaska Judicial Council, Annette Corallo and Vanessa Haney of the Arizona Commission on Judicial Performance Review, Jane Howell of the Colorado Office of Judicial Performance Evaluation, and Joanne Slotnik of the Utah Judicial Performance Evaluation Commission and Tim Shea at the Utah Administrative Office of the Courts. More broadly, we thank the coordinators of JPE programs in other states who participate in our monthly JPE Working Group calls for their willingness to share insights and experiences relevant to this study.

Finally, IAALS thanks New England Law | Boston Professor—and former IAALS Director of Research—Jordan Singer who lent invaluable support to this project both during his time at IAALS and after.

TABLE OF CONTENTS

Executive Summary 1

Introduction..... 3

Overview of Judicial Performance Evaluation..... 4

 Scope & Purpose..... 4

 JPE in the States 5

 Concerns about Judicial Performance Evaluation 6

Data and Methods..... 9

Individual State Results 11

 Alaska 11

 Arizona..... 13

 Colorado 19

 Utah 22

Discussion 25

 Extent of Judicial Experience 25

 Survey Respondent Bias..... 27

Recommendations..... 30

 Comprehensive JPE Programs..... 30

 Well-designed JPE Surveys..... 31

 Broad Education on Implicit Bias..... 32

Conclusion 34

Endnotes 35

Appendix A: State-by-State Overview of JPE Data..... A1

Appendix B: Attorney Survey B1

Appendix C: Court Staff Survey C1

EXECUTIVE SUMMARY

Judicial performance evaluation (JPE) is intended to provide a broad-based and politically neutral assessment of judges' performance on the bench—one that is based on the processes of judging rather than the outcomes reached. Evaluation results may be used to facilitate judicial self-improvement, to inform decisions regarding the retention or reappointment of judges, and to promote public confidence in the judiciary. Particularly with respect to providing evaluation results to voters in the context of judicial retention elections, those who coordinate and who advocate for the adoption of JPE programs must strive to ensure that the evaluations are both fair and accurate.

With the expansion of the use of official JPE programs for retention elections, concerns have been raised about whether women and minority judges are consistently rated lower than their Caucasian male colleagues. A few single-state studies suggest that these concerns may be justified. If this is the case, JPE programs could impede efforts to diversify the bench.

Dedicated to continuous improvement of processes for choosing, evaluating, and retaining judges, IAALS undertook an examination of the four states with longstanding comprehensive and official JPE programs used to assist voters in making retention decisions. In this study, we compare mean scores for women, men, Caucasian, and minority judges for broad performance categories—such as Legal Ability, Impartiality, Communication Skills—and for individual survey questions, to see whether there are consistent differences in evaluation scores based on gender and ethnicity. Our findings are summarized below.

The evaluation results examined here confirm the high quality of the justices and judges serving in these four states.

For Alaska judges, average scores for broad performance categories range from 3.7 to 4.9, on a 1-5 scale. In Arizona, overall evaluation scores range from a low of 2.9 to a high of 3.8, on a 0-4 scale. Grades for Colorado judges range from a B (a 3.1 average) to a B+/A- (a 3.6 average). In Utah, across all evaluation items, at least 87 percent of respondents give judges a favorable rating.

Most of the differences we find between evaluation scores for women and men judges, and for minority and Caucasian judges, are small—typically no more than one-tenth of a point or no larger than three percent, depending on the scale of measurement used.

In Alaska, attorneys rated women judges two-tenths of a point lower than men judges on Impartiality/Fairness and Temperament. Arizona attorneys rated minority trial judges lower than Caucasian judges on Legal Ability (three-tenths of a point lower) and Communication Skills (two-tenths of a point lower). In Colorado, minority trial judges scored two-tenths of a point lower than Caucasian judges in attorney assessments of their willingness to reconsider errors and to handle complicated and time-consuming cases. In Utah, the largest disparities were found in attorney evaluations of women and men judges with respect to avoiding ex parte communications, with men rated “Excellent” five percent more often than women.

The largest differences in evaluation scores based on gender and ethnicity are in areas where past research suggests that implicit biases may come into play.

The gender- and ethnicity-based differences found in evaluation scores cannot be definitively attributed to one factor or another, but past research suggests that implicit biases may provide one potential explanation—that those who evaluate judges may unconsciously rely on stereotypes or fixed notions about appropriate roles and behaviors for women and men and for minorities and non-minorities. Respondents may be invoking these biases in assessing such qualities as judicial competence and judicial demeanor.

Steps may be taken to limit the effects of potential implicit biases on judicial performance evaluations.

If implicit biases are a factor, then it falls to proponents of JPE to work to address them. We offer three recommendations that can serve that objective. (1) State JPE programs should be broad-based, in terms of the types of respondents who are surveyed and the supplemental evaluation tools that are used. It is particularly important that respondent groups in addition to attorneys be included. (2) Evaluation surveys should be developed in consultation with experts in the field of job-performance evaluation and survey design, and should focus respondents' attention on observable behaviors, both positive and negative. (3) Bar associations and court systems should take steps to raise awareness of, and promote education about, the potential for implicit biases to influence thoughts and decisions.

INTRODUCTION

IAALS, the Institute for the Advancement of the American Legal System at the University of Denver, has worked in the area of judicial performance evaluation (JPE) since opening its doors in January 2006. Dedicated to continuous improvement in all areas in which it works, IAALS seeks to assist jurisdictions in developing empirically based systems for judicial performance evaluation that preserve impartiality and provide accountability.

An official evaluation program—i.e., one that is authorized by constitution, statute, or court rule—is in place in 18 states and the District of Columbia. In eight states, JPE results are used, at least in part, to provide objective, broad-based, and apolitical information to voters for judicial retention elections. Several additional jurisdictions are considering the establishment of JPE programs for this purpose, as well as for the purpose of providing feedback to judges. At the same time, concerns have been raised that judicial performance evaluation disadvantages women and minority judges, who are perceived as scoring systematically lower in evaluations than their Caucasian and male counterparts. Critics suggest that stereotypes about women and minorities find their way into assessments of the work of these judges. A few single-state studies suggest that these concerns may be valid. If this is the case, JPE programs may be an impediment to efforts to diversify the bench.

Dedicated to continuous improvement of processes for choosing, evaluating, and retaining judges, IAALS undertook a comprehensive examination of JPE programs in four of the eight states where results are used, at least in part, to assist voters in making retention decisions.

“ JPE IS BASED ON MORE THAN JUST POLITICALLY NEUTRAL EXPECTATIONS; IT IS BASED ON GENDER-NEUTRAL AND ETHNICALLY NEUTRAL EXPECTATIONS AS WELL. ”

OVERVIEW OF JUDICIAL PERFORMANCE EVALUATION

SCOPE & PURPOSE

Judicial performance evaluation is a mechanism through which judges' performance on the bench can be measured against established (and apolitical) benchmarks. Although JPE programs vary in their scope and design, most consist of the collection and analysis of behavior- and process-oriented data on judicial performance in such areas as legal ability, impartiality and fairness, integrity, judicial temperament and demeanor, communication skills, administrative skills, and professionalism. The most common tools used in these programs are surveys of those who interact with the judge professionally. Among potential respondents, attorneys are the most frequently surveyed group due to their recurring interactions with judges. Other respondent groups surveyed as part of the JPE process are non-attorneys who frequently enter the courtroom (e.g., jurors, litigants, police and probation officers), court staff, and other judges. Some JPE commissions also review judicial decisions and opinions (especially relevant for evaluations of appellate judges) and case management statistics, conduct interviews of the evaluated judge, consider self-evaluations, and/or incorporate courtroom observation.

JPE serves a number of important functions that have been recognized by a broad cross-section of stakeholders. First, all JPE programs provide an opportunity for professional development. Similar to routine performance evaluations undertaken in almost every workplace, JPE gives judges feedback that enables and encourages self-improvement. Because attorneys and others who frequently interact with judges have little opportunity to provide meaningful and honest feedback on performance, JPE offers constructive feedback that is otherwise unavailable. A 2008 study by IAALS, *The Bench Speaks on Judicial Performance: A Survey of Colorado Judges*, found that “judges largely believed that the survey data—especially from attorneys—was valuable to their professional development.”¹ One respondent noted that “as a judicial officer feedback either positive or negative is not readily available [outside the JPE process] and if received it is generally result oriented and therefore suspect.”² Another commented that “[t]he ability to receive feedback from the public is a wonderful way to assess my performance. Even the critical comments are helpful.”³ More broadly, judges surveyed indicated that going through the JPE process was beneficial to their professional development.⁴ This was especially the case with respect to trial judges, nearly one-fifth of whom said JPE was “significantly beneficial” and 66 percent of whom said it was “somewhat beneficial.”⁵

JPE is also a valuable resource for those charged with deciding whether judges should be retained for additional terms. For voters in states with retention elections, JPE offers substantive information on which to base decisions. Studies suggest that when lacking information about judges, voters may make decisions based on criteria such as ethnicity or gender.⁶ A 2007 poll of Colorado voters conducted by IAALS and the League of Women Voters of Colorado delved into voting strategies of the “uninformed voter” and discovered the following: 28 percent of respondents vote to keep all judges; 16 percent vote to remove all judges; 24 percent vote randomly for/against judges; seven percent vote on surname; and three percent vote on gender.⁷ Other studies suggest that where voters lack information about judges, they may vote based on such factors as party affiliation, name recognition, or ballot position, or decline to vote entirely.⁸ A recent article by Professor Jordan Singer suggests that voters prefer to choose judges based on procedural fairness criteria, although they will turn back to partisan messaging if information on procedural fairness is unavailable.⁹ The information provided by JPE ensures that voters have adequate information to allow them to cast an informed vote, as well as that the information they have is neutral and performance- rather than issue-based. Performance evaluations serve a similar informative function in states where judicial reappointment rests with the legislature.

By providing relevant information to voters and others, JPE ensures that judges are held appropriately accountable for their work without infringing upon their decisional independence. In fact, JPE has the potential to enhance judicial independence. By focusing on neutral, nonpartisan, process-oriented standards, rather than on particular case outcomes, JPE ensures that judges can follow the rule of law while being held to broadly accepted standards for their overall performance. Of the Colorado judges surveyed by IAALS, 24 and 18 percent of appellate and trial judges, respectively, indicated that JPE “moderately increases” judicial independence, with one survey respondent noting that “[t]he process itself increases judicial independence by assuring the public that judges are subject to a review in which voters play a part.”¹⁰ A plurality of both appellate and trial judges (41 and 44 percent, respectively) indicated that JPE had “no effect” on their judicial independence.¹¹

Finally, JPE promotes public trust and confidence in the courts. As noted in a 2007 *Judicature* article, “[p]roviding the results of individual judicial evaluations to the electorate ... in a manner that is easily understood builds trust and confidence in the judiciary by identifying judges with outstanding performance and identifying those who need improvement.”¹² Members of the public will place greater faith in an institution whose officers are held to, and assessed on, job performance standards—just as they are in their own work.

JPE IN THE STATES

The official JPE programs that operate in varying forms in 18 states and the District of Columbia differ considerably.¹³ In eight states, JPE results are provided to voters for use in retention elections: Alaska, Arizona, Colorado, Kansas, Missouri, New Mexico, Tennessee, and Utah. The programs are administered by a performance evaluation commission established for that purpose, usually comprised of both attorneys and non-attorneys, or by an existing entity within the state administrative office of the courts such as a judicial council. These bodies oversee the collection of performance data, produce evaluation reports for voters, and ensure broad dissemination.

In three states—Connecticut, New Jersey, and Vermont—and the District of Columbia, results of judicial performance evaluations are provided to those responsible for reappointing judges, but evaluation results are not available to the public. In Hawaii and New Hampshire, summary results are released to the public to enhance public trust and confidence in the courts. These narratives do not identify individual judges, but rather report aggregate data on the performance of the court as a whole. Finally, in five states—Florida, Idaho, Illinois, Massachusetts, and Rhode Island—results are shared only with individual judges.

Some of the aforementioned programs apply to all judges in the state (i.e., trial and appellate); others apply only to judges of particular courts or judges subject to a particular method of selection and retention. Similarly, there is considerable variation from state to state in the procedures used and the data collected for evaluations. The more comprehensive evaluation programs employ a variety of assessment tools, including attorney surveys, non-attorney surveys, court staff surveys, trial and/or appellate judge surveys, peer judge surveys, self-evaluation, courtroom observation, data on recusal and reversal rates, disciplinary records, written opinion review, public comment, interviews with the evaluated judge, and caseload management statistics. With respect to reporting the results to the public (in those states that do so), full data reports are posted online, along with narrative summaries of evaluation results.

Unofficial JPE programs also exist in a number of states. For example, in North Carolina, the North Carolina Bar Association recently launched an evaluation program both for incumbent judges seeking reelection and for their attorney challengers.¹⁴ Bar associations in other states also offer judicial polls, plebiscites, and ratings, and in one state, a newspaper evaluates supreme court justices and trial court judges in one county.¹⁵ These programs differ from the official programs described above in many respects, but the most fundamental difference is that the evaluations are based only on surveys of attorneys.

CONCERNS ABOUT JUDICIAL PERFORMANCE EVALUATION

JPE is intended to provide a politically neutral measure of judges' ability and performance on the bench, based not on substantive outcomes but rather on established benchmarks for the processes that led to those outcomes. Thus, voters and others are able to make retention and reappointment decisions based on these broadly applicable benchmarks rather than on factors such as party affiliation, ethnicity, and gender. In this context, JPE is based on more than just politically neutral expectations; it is based on gender-neutral and ethnically neutral expectations as well. This is particularly important in promoting and ensuring diversity in the judiciary—a profession that does not and traditionally has not reflected the increasing diversity of our society. Consider the following:

Today, white males are overrepresented on state appellate benches by a margin of nearly two-to-one. Almost every other demographic group is underrepresented when compared to their share of the nation's population. There is also evidence that the number of black male judges is actually decreasing There are still fewer female judges than male, despite the fact that the majority of today's law students are female, as are approximately half of all recent law degree recipients. This pattern is most prevalent in states' highest courts, where women have historically been almost completely absent.¹⁶

Our report does not comment on the merits of various selection systems, although it is worth noting that “both elective and appointive systems are producing similarly poor outcomes in terms of the diversity of judges.”¹⁷ To be sure, efforts to increase diversity on the bench must be focused on the selection stage in the first instance, but JPE also has the potential to play an important role in this effort. JPE programs ensure that, once on the bench, judges are retained (or not retained) on the basis of their performance on the bench, not on the basis of their gender or ethnicity.

JPE has the capacity, therefore, to promote judicial diversity by providing decision makers with gender- and ethnically neutral information on judicial performance. Concerns have been raised, however, that respondents systematically score female/minority judges lower than male/Caucasian judges in JPE surveys, disadvantaging those judges in the eyes of those responsible for retaining/reappointing them and hindering diversity on the bench. Of particular concern are the subjectively based tools used in the evaluation process. Although a few states look to court and caseload data, recusal rates, and reversal rates, performance evaluation programs invariably require some degree of subjective evaluation on the part of the respondents/observers. It is through these subjective tools, some argue, that biases against women and minorities are translated into disproportionately low evaluation scores for these groups. Over the past few decades, a number of discrete reviews have been undertaken and commentaries published on these perceived discrepancies between judicial performance evaluation scores of female and minority judges, and male and Caucasian judges.

Following up on a 1990 Colorado Supreme Court study on gender bias, University of Denver Law Professor Joyce Sterling examined whether there was evidence of gender bias in attorney perceptions of judges by analyzing the 1992 Colorado judicial performance evaluation results.¹⁸ Professor Sterling found “a clear pattern of bias that emerges” from the data, with respect to both overall performance and specific attributes.¹⁹ With respect to overall performance (measured by the following survey question: “What is your overall evaluation of this judge?”), the data showed that almost 25 percent of attorneys recommended that female judges not be retained, while only 13 percent recommended that male judges not be retained.²⁰ When data for specific attributes was examined, “it was apparent that female judges were rated lower consistently than their male counterparts on every attribute measured” and “[i]n every instance, these differences were statistically significant.”²¹ Although there was more variation in the responses of female attorneys than in those of male attorneys, there were select areas in which female attorney respondents ranked female judges significantly lower than male judges, particularly in the areas of compassion, courtesy, satisfactory performance as a motions judge, satisfactory performance as a settlement judge, and overall rating.²² According to Professor Sterling, the “particular character of the ratings suggests that evaluations may be influenced by gender stereotypes.”²³

In 1987, the American Bar Association (ABA) created the Commission on Women in the Profession, whose mission is “to secure full and equal participation of women in the ABA, the profession, and the justice system.”²⁴ The Commission examined experiences of women in various sectors of the legal profession and in its December 1995 report *Unfinished Business: Overcoming the Sisyphus Factor* concludes that “even women who enjoy the prestige of the judiciary are affected by bias.”²⁵ The report asserts the following with respect to judicial performance evaluations:

[W]omen judges endure consistently stronger criticism than their male colleagues, especially in subjective categories such as demeanor. Those who score high in response to judicial evaluation questions about legal knowledge, promptness and case management are subject to condescending barbs. Women judges often receive low marks for strong and decisive action, behavior that garners praise for their male colleagues.²⁶

The Commission recommended that JPE programs “closely scrutinize subjective criticisms to determine whether gender bias is the underlying motivation.”²⁷

In response to increasing use of judicial performance evaluation, Utah Supreme Court Justice Christine M. Durham (later Chief Justice Durham) in 2000 explored the issue of bias in judicial performance evaluation programs. Citing to general information compiled by various state committees on gender bias in the courts and to Professor Sterling’s study of the 1992 Colorado judicial performance evaluation surveys of attorneys, Justice Durham preliminarily concludes that “there are widespread perceptions and some preliminary empirical findings suggesting that women judges, certainly the most powerful women in the courts, are not insulated from the effects of gender bias.”²⁸ She also points to the probability that the judicial performance evaluation process is subject to gender-related bias, based on support from social science and psychological literature showing that women in positions of power encounter a “role conflict” phenomenon.²⁹ Recognizing the largely implicit and indirect nature in which expectations and assumptions about gender-related characteristics operate, Justice Durham challenges the extent to which performance evaluation programs “permit or encourage the exercise of subjective judgment.”³⁰

Some have voiced concerns that biases are affecting retention decisions in Hawaii. A 2004 article in the *Honolulu Star-Bulletin* called attention to the fact that five of the six state judges not reappointed since 2001 were women.³¹ Commentators also noticed a similar trend with respect to Hawaii State Bar Association ratings of judicial nominees. In the years just prior to the 2004 article, nine of 11 male nominees were rated highly qualified by the Bar while only one of six female nominees was rated highly qualified. One of the female nominees, a deputy prosecutor, was applying for a seat on the Family Court, and the Bar rated her unqualified for this position due to her alleged lack of Family Court experience. By contrast, the Bar rated a male nominee (and a former court administrator) as highly qualified for a Family Court position despite his lack of Family Court experience. Then-Hawai'i Senator Colleen Hanabusa said of these perceived trends: "I find it difficult to accept that women who have been nominated as judges are consistently not as qualified and the women judges up for retention are lesser qualified than the men. Something is going on."³² Until very recently, the Bar was not required to provide any explanation to accompany nominee ratings. In the aftermath of a controversy over the Bar's 2010 negative ("unqualified") rating of Katherine Leonard, a chief justice nominee, the policy was changed to allow the Bar president to disclose the reasons behind the ratings with the approval of the board of directors.³³

In 2007, University of Missouri-St. Louis Professor Gary K. Burger released a statistical analysis of The Judicial Evaluation Survey of attorneys, conducted by the Missouri Bar, the Bar Association of Metropolitan St. Louis, and the Kansas City Metropolitan Bar Association for the purpose of providing voters with a clearer picture of judges' qualifications. This study was undertaken in response to concerns that, without safeguards that ensure that only attorneys with "personal knowledge" respond to the surveys, there was too great a chance that "a lawyer or group of lawyers at odds with a judge over an opinion or ruling, or for any other reason, [could] manipulate the results."³⁴ Using the data that was published in voter information pamphlets for five prior evaluation cycles (1998, 2000, 2002, 2004, and 2006), the study focused on differences associated with the type of court, gender, and ethnicity of the particular judge as each related to the percentage of favorable retention recommendations. The data showed that males were recommended for retention at a significantly higher rate (85.84 percent) than were females (76.69 percent), and that Caucasians were recommended for retention at a significantly higher rate (85.44 percent) than were African Americans (75.19 percent).³⁵ These trends were found in all three of the courts studied.³⁶ Looking at gender and ethnicity combinations, the analysis showed that African American females were recommended for retention at a significantly lower rate (66.37 percent) than Caucasian females (83.10 percent), African American males (84.02 percent), and Caucasian males (86.21 percent).³⁷ African American females were also rated significantly lower in the area of "legal analysis."³⁸ Possible explanations include real differences in performance, problems with rating scales, problems in the evaluation process, bias on the part of respondents, or a combination of these factors.³⁹

Most recently, UNLV Professors Rebecca Gill and Sylvia Lazos, along with graduate student Mallory Waters, analyzed the bar poll administered biennially by the *Las Vegas Review Journal* to rate Clark County judges and supreme court justices. Looking at ten years of data (1998 to 2008), Gill, Lazos, and Waters report a "large, unexplained gap in the ratings of female and minority judges and their male and nonminority counterparts, all other measures of judicial quality being equal."⁴⁰ As "objective measures of performance quality," the authors used reversal rates, selection method (election or interim appointment), legal education, judicial experience, and ethical record.⁴¹ When these objective measures are held constant, the study showed that retention scores for women judges are 11 points lower than those of their male counterparts, and scores for minority judges are 14 points lower than for their Caucasian counterparts.⁴² Although cautioning that "the evidence is not good for the LVRJ's privately run JPE," the authors suggest a need for "further inquiry and innovation."⁴³

DATA AND METHODS

In light of these concerns, IAALS undertook a study of official JPE programs used to inform voters in judicial retention elections, in order to assess whether women and/or minority judges consistently receive lower scores than their Caucasian male colleagues. IAALS limited this study to the four states with longstanding comprehensive and official (i.e., authorized by constitution, statute, or court rule) retention evaluation programs: Alaska, Arizona, Colorado, and Utah.⁴⁴ Working in close collaboration with the JPE program coordinators in these states, IAALS compiled evaluation data from JPE surveys over the course of several evaluation cycles and ascertained the gender, ethnicity, and extent of judicial experience for each evaluated judge.

- Alaska: data for the 2006, 2008, and 2010 evaluation cycles
- Arizona: data for the 2002, 2004, 2006, 2008, and 2010 evaluation cycles
- Colorado: data for the 2004, 2006, 2008, and 2010 evaluation cycles
- Utah: data for the 2002, 2004, 2006, 2008, and 2010 evaluation cycles

There is variation across the evaluation programs in these four states in terms of the evaluative criteria, survey questions, respondent groups, measurement, and scoring, so we examine each state individually. (See Appendix A for a state-by-state overview of the JPE data used in this study.) For each state, we endeavored to include as many retention cycles as possible. We analyzed each state program as it operated in 2010 (and as described in the program descriptions that follow), and we went backward in time to add as many retention cycles of data as was feasible, based on consistency in questions asked and respondents surveyed.⁴⁵ As such, the years examined for each state vary.

To assess whether there are meaningful differences in how the performance of female and minority judges is evaluated, we compare mean scores for performance categories (e.g., Legal Ability, Impartiality, Communication Skills) and for individual survey questions. We confine these comparisons to groups (e.g., women judges, Caucasian judges) that include at least ten judges, to moderate the influence of outliers. We discuss our findings below, and we highlight instances where evaluation scores for men and Caucasian judges were notably higher than scores for women and minority judges.⁴⁶

To illustrate our approach, we begin by comparing scores for women and minority judges to those of men and Caucasian judges in the one area that is consistent across all four state programs—attorney evaluations of judges' overall performance. As does each of the figures that follow, **Figure 1** (pg. 10) displays mean scores, and differences in mean scores, for women and men judges and minority and Caucasian judges. A positive difference in mean scores indicates that female judges scored higher than male judges, or that minority judges scored higher than Caucasian judges. A negative difference indicates the opposite—that female and minority judges scored lower than their male and Caucasian colleagues.

FIGURE 1 :: Attorney Evaluations of Judges

	WOMEN	MEN	DIFFERENCE	MINORITY	CAUCASIAN	DIFFERENCE
Alaska (2006 - 2010) : Average (1 - 5)	4.0	4.1	-0.1	--	--	--
Arizona (2002 - 2010) : Average (0 - 4)	3.2	3.3	-0.1	3.2	3.3	-0.1
Colorado (2004 - 2010) : Average (0 - 4)	3.3	3.4	-0.1	3.3	3.4	-0.1
Utah (2002 - 2010) : % Favorable	93%	94%	-1%	93%	94%	-1%

In all four states for the retention cycles examined, attorneys gave higher overall evaluations to male and Caucasian judges. However, in each instance, differences in scores based on gender and ethnicity are small. For Alaska, Arizona, and Colorado, where judges are assessed on a five-point scale, differences are no greater than one-tenth of a point. For Utah, where percentages of attorneys rating judges favorably are reported, differences are no greater than one percent. We do not include a comparison of minority and Caucasian judges for Alaska because only one minority judge (i.e., fewer than our minimum of ten judges) stood for retention from 2006 to 2010.

“ [T]HE RESULTS DISCUSSED IN THIS REPORT CONFIRM THE HIGH QUALITY OF THE JUSTICES AND JUDGES SERVING IN THESE FOUR STATES. ”

INDIVIDUAL STATE RESULTS

ALASKA

As the first state to establish an official JPE program, Alaska has evaluated all justices and judges prior to their retention elections since 1976. Here we analyze evaluation results from 2006 to 2010.

For supreme court justices and court of appeals judges, the Alaska Judicial Council surveys all active and inactive members of the Alaska Bar Association and court employees. For superior court and district court judges, surveys are sent to five respondent groups: attorneys; police and probation officers who handle criminal cases; court employees; jurors; and social workers, guardians ad litem, and child advocates. Since 2006, those who evaluate justices and judges rate them on a five-point scale—Poor (1), Deficient (2), Acceptable (3), Good (4), Excellent (5)—in five areas, including legal ability (for attorneys only), impartiality, integrity, temperament, and diligence.

In Alaska, as in the other three states, we combined data from different court levels where the questions asked of each respondent group were the same, in order to capitalize on the available data. Here we combine evaluation data for judges of all courts—supreme court, court of appeals, superior court, and district court.

Figure 2 (pg. 12) displays mean scores, and differences in mean scores, for women and men judges.⁴⁷ Both groups fared well, with average scores for women ranging from 3.7 to 4.9 and averages for men ranging from 4.0 and 4.9.

For justices and judges evaluated from 2006 to 2010, men scored better in evaluations by attorneys and by police and probation officers, though the differences are two-tenths of a point or smaller. For attorneys, the greater disparities between scores for women and men judges are found for Impartiality/Fairness and Temperament; for police and probation officers, the wider gaps are found for Impartiality/Fairness, Temperament, Diligence, and in the overall assessment.

At the same time, women received higher scores from court employees—one-tenth of a point higher on each criteria, and from social workers—at least two-tenths of a point higher for Temperament, Diligence, and the overall evaluation. There were no differences in juror assessments of the performance of women and men judges.

FIGURE 2 :: ALASKA :: Evaluations of Appellate and Trial Judges

	AVERAGE (1 – 5)		
	WOMEN	MEN	DIFFERENCE
ATTORNEYS^a			
Legal Ability	4.0	4.0	0.0
Impartiality/ Fairness	3.9	4.1	-0.2
Integrity	4.2	4.3	-0.1
Temperament	3.9	4.1	-0.2
Diligence	4.1	4.1	0.0
OVERALL	4.0	4.1	-0.1
COURT EMPLOYEES^a			
Impartiality/ Fairness	4.4	4.4	0.1
Integrity	4.5	4.4	0.1
Temperament	4.4	4.3	0.1
Diligence	4.4	4.4	0.1
OVERALL	4.4	4.4	0.1
POLICE AND PROBATION OFFICERS^b			
Impartiality/ Fairness	3.7	4.0	-0.2
Integrity	4.0	4.2	-0.1
Temperament	3.9	4.1	-0.2
Diligence	3.9	4.1	-0.2
OVERALL	3.8	4.0	-0.2
SOCIAL WORKERS^c			
Impartiality/ Fairness	4.3	4.4	-0.1
Integrity	4.5	4.5	0.0
Temperament	4.5	4.3	0.2
Diligence	4.6	4.3	0.3
OVERALL	4.7	4.4	0.3
JURORS^d			
Fair and impartial to all sides	4.8	4.8	0.0
Respectful and courteous to parties	4.9	4.9	0.0
Attentive during proceedings	4.8	4.8	0.0
Exercised control over proceedings	4.8	4.8	0.0
Intelligence and skill as a judge	4.8	4.8	0.0
OVERALL PERFORMANCE	4.8	4.8	0.0

^a N(Women)=13; N(Men)=58.

^b Trial judges only. N(Women)=12; N(Men)=55.

^c Trial judges only. Includes social workers, guardians ad litem, and court-appointed special advocates. N(Women)=11; N(Men)=44.

^d Trial judges only. N(Women)=12; N(Men)=55.

ARIZONA

Arizona voters approved a constitutional amendment creating a judicial performance evaluation program in 1992, making Arizona the only state with a constitutionally authorized program. The program is limited to judges subject to the state's merit selection and retention system, which includes all appellate justices and judges and trial judges in the three largest counties—Maricopa, Pima, and Pinal.⁴⁸ (Pinal County did not move to merit selection and retention for superior court judges until 2012, so no evaluation results for these judges are included here.) Judges are evaluated midterm and prior to standing for retention. Here we analyze retention evaluation results from 2002 to 2010.

For supreme court justices and court of appeals judges, the Commission on Judicial Performance Review surveys attorneys (including staff attorneys), peer justices and judges, and superior court judges. For superior court judges, surveys are sent to attorneys, non-attorneys (including litigants, witnesses, and self-represented litigants), jurors, and court staff. Evaluations by peer judges and court staff are confidential.

For all justices and judges, respondents are asked a variety of questions related to legal ability (for attorneys and judges only), integrity, communication skills, judicial temperament, and administrative performance. For trial court judges, respondents are also asked questions relating to judges' settlement activities. On each question, justices and judges are rated on a five-point scale—Unacceptable (0), Poor (1), Satisfactory (2), Very Good (3), Superior (4). Average scores are reported.

As shown in **Figure 3** (pg. 14), attorney evaluations of appellate judges are strong, with categorical scores for both women and men judges ranging from 3.0 to 3.5 on a 0-4 scale.⁴⁹ Differences in mean scores for women and men judges are no greater than one-tenth of a point. Men scored slightly higher than women on Legal Ability and on three of the four questions in this area, but women scored higher on "Demeanor in communications with counsel" and "Promptness in making rulings and rendering decisions."

Appellate judges also fared well in evaluations by trial judges, as seen in **Figure 4** (pg. 15). Scores for performance categories ranged from 3.3 to 3.7. However, we see more differences between scores for women and men judges in trial judge evaluations than in attorney evaluations. Women appellate judges scored one-tenth of a point lower than their male counterparts on Legal Ability and Administrative Performance and on eight of 14 individual questions, including questions related to Integrity.

As **Figure 5** (pg. 16) shows, attorney evaluations of trial judges fall in a lower range than those of appellate judges. Minority judges received the lowest average score for all groups—2.9 on a 0-4 scale—on Legal Ability. Women, men, and Caucasian judges scored a high of 3.4 on Integrity.

FIGURE 3 :: ARIZONA :: Attorney Evaluations of Appellate Judges

	AVERAGE (0 – 4)		
	WOMEN	MEN	DIFFERENCE
LEGAL ABILITY	3.0	3.1	-0.1
Legal reasoning ability	3.1	3.1	-0.1
Knowledge of the law	3.1	3.1	-0.1
Decisions based on law and facts	3.0	3.0	0.0
Clearly written, legally supported decisions	3.0	3.1	-0.1
INTEGRITY	3.5	3.5	0.0
Equal treatment regardless of race	3.5	3.5	0.0
Equal treatment regardless of gender	3.5	3.5	0.0
Equal treatment regardless of religion	3.6	3.5	0.0
Equal treatment regardless of national origin	3.5	3.5	0.0
Equal treatment regardless of disability	3.6	3.5	0.0
Equal treatment regardless of age	3.6	3.5	0.0
Equal treatment regardless of sexual orientation	3.5	3.5	0.0
Equal treatment regardless of economic status	3.4	3.4	0.0
Basic fairness and impartiality	3.4	3.4	0.0
COMMUNICATION SKILLS	3.4	3.4	0.0
Attentiveness	3.4	3.4	0.0
Demeanor in communications with counsel	3.5	3.4	0.1
Relevant questions	3.3	3.3	0.0
Preparation for oral argument	3.4	3.4	0.0
JUDICIAL TEMPERAMENT	3.4	3.4	0.0
Dignified	3.5	3.5	0.0
Courteous	3.5	3.5	0.0
Patience	3.4	3.4	0.0
Conduct that promoted public confidence in the court and the judge's ability	3.4	3.4	0.0
ADMINISTRATIVE PERFORMANCE	3.2	3.1	0.1
Promptness in making rulings and rendering decisions	3.2	3.1	0.1

N(Women)=10; N(Men)=31.

FIGURE 4 :: ARIZONA :: Trial Judge Evaluations of Appellate Judges

	AVERAGE (0 – 4)		
	WOMEN	MEN	DIFFERENCE
LEGAL ABILITY	3.3	3.4	-0.1
Legal reasoning ability	3.3	3.4	0.0
Knowledge of the law	3.3	3.4	-0.1
Decisions based on law and facts	3.3	3.4	-0.1
Clearly written, legally supported decisions	3.3	3.4	-0.1
INTEGRITY	3.6	3.7	0.0
Equal treatment regardless of race	3.6	3.7	0.0
Equal treatment regardless of gender	3.6	3.7	0.0
Equal treatment regardless of religion	3.6	3.7	-0.1
Equal treatment regardless of national origin	3.6	3.7	0.0
Equal treatment regardless of disability	3.6	3.7	-0.1
Equal treatment regardless of age	3.6	3.7	-0.1
Equal treatment regardless of sexual orientation	3.6	3.7	0.0
Equal treatment regardless of economic status	3.6	3.6	-0.1
Basic fairness and impartiality	3.7	3.7	0.0
ADMINISTRATIVE PERFORMANCE	3.3	3.4	-0.1
Promptness in making rulings and rendering decisions	3.3	3.4	-0.1

N(Women)=10; N(Men)=31.

Differences in scores between women and men judges are no greater than one-tenth of a point, but disparities between minority and Caucasian judges are as large as three-tenths of a point for Legal Ability and two-tenths of a point for Communication Skills.

As seen in **Figure 6** (pg. 17), judges in all groups fared better among non-attorneys. The lowest average score (3.2) was for women and minority judges on Communication Skills, while the highest average score (3.5) was for women, men, and Caucasian judges on Integrity and for Caucasian judges on Judicial Temperament.

While differences in non-attorney evaluation scores between women and men judges are no greater than one-tenth of a point, minority judges scored two-tenths of a point lower than their Caucasian counterparts on four of 19 questions, including three items related to Judicial Temperament.

Figure 7 (pg. 18) offers evidence of a so-called “halo effect” in juror evaluations of judges, where jurors tend to rate judges higher than other respondent groups.⁵⁰ The highest averages (3.8) are found for minority judges on Communication Skills and for women and minority judges on Judicial Temperament. The “lowest” averages of 3.6 are found for all groups on Administrative Performance.

With the exception of a single survey item, there are no differences in juror evaluations of judges based on gender or ethnicity. Both women and minority judges scored one-tenth of a point lower than their Caucasian male counterparts on “Was prepared for the proceedings.”

FIGURE 5 :: ARIZONA :: Attorney Evaluations of Trial Judges

	AVERAGE (0 – 4)					
	WOMEN	MEN		MINORITY	CAUCASIAN	DIFFERENCE
LEGAL ABILITY	3.0	3.1	-0.1	2.9	3.2	-0.3
Legal reasoning ability	3.0	3.1	-0.1	2.9	3.1	-0.3
Knowledge of substantive law	3.0	3.1	-0.1	2.9	3.1	-0.3
Knowledge of rules of evidence	3.5	3.2	-0.1	2.9	3.2	-0.3
Knowledge of rules of procedure	3.1	3.2	-0.1	2.9	3.2	-0.2
INTEGRITY	3.4	3.4	0.0	3.3	3.4	-0.1
Basic fairness and impartiality	3.2	3.3	-0.1	3.1	3.3	-0.1
Equal treatment regardless of race	3.4	3.4	0.0	3.4	3.4	-0.1
Equal treatment regardless of gender	3.4	3.4	-0.1	3.3	3.4	-0.1
Equal treatment regardless of religion	3.4	3.5	0.0	3.4	3.5	-0.1
Equal treatment regardless of national origin	3.4	3.4	0.0	3.4	3.5	-0.1
Equal treatment regardless of disability	3.4	3.5	0.0	3.4	3.5	-0.1
Equal treatment regardless of age	3.4	3.4	0.0	3.4	3.5	-0.1
Equal treatment regardless of sexual orientation	3.4	3.5	0.0	3.4	3.5	-0.1
Equal treatment regardless of economic status	3.4	3.4	0.0	3.3	3.4	-0.1
COMMUNICATION SKILLS	3.1	3.2	0.0	3.0	3.2	-0.2
Clear and logical oral communications and directions	3.1	3.2	0.0	3.0	3.2	-0.2
Clear and logical written decisions	3.1	3.1	0.0	3.0	3.1	-0.2
Gave all parties an adequate opportunity to be heard ^a	3.2	3.2	0.0	3.2	3.3	-0.1
JUDICIAL TEMPERAMENT	3.2	3.2	-0.1	3.2	3.2	-0.1
Understanding and compassion	3.1	3.2	-0.1	3.1	3.2	-0.1
Dignified	3.2	3.3	-0.1	3.2	3.3	-0.1
Courteous	3.2	3.3	-0.1	3.2	3.3	-0.1
Conduct that promotes public confidence in the court and judges' ability	3.2	3.2	-0.1	3.1	3.2	-0.1
Patient ^a	3.2	3.2	0.0	3.2	3.2	0.0
ADMINISTRATIVE PERFORMANCE	3.2	3.2	-0.1	3.1	3.2	-0.1
Punctual(ity) in conducting proceedings	3.2	3.2	0.0	3.1	3.2	-0.1
Maintained (maintenance of) proper control of (over) courtroom	3.2	3.3	-0.1	3.2	3.3	-0.1
Prompt(ness) in making rulings and rendering decisions	3.2	3.2	-0.1	3.0	3.2	-0.2
Was prepared for the proceedings ^a	3.3	3.3	0.0	3.2	3.3	-0.1
Efficient management of calendar	3.1	3.2	-0.1	3.1	3.2	-0.1
SETTLEMENT ACTIVITIES	3.2	3.2	0.0	3.1	3.2	-0.1
Appropriately promoted or conducted settlement/ Appropriate actions in encouraging settlement negotiations	3.2	3.2	0.0	3.1	3.2	-0.1

N(Women)=87; N(Men)=200; N(Minority)=58; N(Caucasian)=229.

^aThese questions were asked in 2008 and 2010 only, so the number of evaluated judges in each category is smaller: N(Women)=38; N(Men)=77; N(Minority)=23; N(Caucasian)=92.

FIGURE 6 :: ARIZONA :: Non-Attorney Evaluations of Trial Judges

	AVERAGE (0 – 4)					
	WOMEN	MEN	DIFFERENCE	MINORITY	CAUCASIAN	DIFFERENCE
INTEGRITY	3.5	3.5	0.0	3.4	3.5	-0.1
Basic fairness and impartiality	3.3	3.4	-0.1	3.3	3.4	-0.2
Equal treatment regardless of race	3.5	3.5	0.0	3.4	3.5	-0.1
Equal treatment regardless of gender	3.4	3.5	-0.1	3.4	3.5	-0.1
Equal treatment regardless of religion	3.5	3.5	0.0	3.5	3.5	0.0
Equal treatment regardless of national origin	3.5	3.5	0.0	3.4	3.5	-0.1
Equal treatment regardless of disability	3.5	3.5	0.0	3.4	3.5	-0.1
Equal treatment regardless of age	3.5	3.5	0.0	3.4	3.5	-0.1
Equal treatment regardless of sexual orientation	3.5	3.5	0.0	3.5	3.5	0.0
Equal treatment regardless of economic status	3.4	3.5	-0.1	3.4	3.5	-0.1
COMMUNICATION SKILLS	3.2	3.3	0.0	3.2	3.3	-0.1
Explained proceedings (to the jury) ^a	3.3	3.3	-0.1	3.3	3.3	-0.1
Explained reasons for delays ^a	3.2	3.2	-0.1	3.1	3.2	-0.1
JUDICIAL TEMPERAMENT	3.4	3.4	-0.1	3.3	3.5	-0.2
Understanding and compassion	3.3	3.4	-0.1	3.2	3.4	-0.2
Dignified	3.5	3.5	0.0	3.4	3.5	-0.1
Courteous	3.4	3.5	-0.1	3.4	3.5	-0.1
Conduct that promotes public confidence in the court and judges' ability	3.4	3.4	0.0	3.3	3.5	-0.2
Patient ^a	3.3	3.4	-0.1	3.2	3.4	-0.2
ADMINISTRATIVE PERFORMANCE	3.4	3.4	0.0	3.3	3.4	-0.1
Punctual(ity) in conducting proceedings	3.3	3.3	0.0	3.2	3.3	-0.1
Maintained (maintenance of) proper control of (over) courtroom	3.5	3.5	0.0	3.4	3.5	-0.1
Was prepared for the proceedings	3.4	3.4	0.0	3.3	3.4	-0.1

N(Women)=77; N(Men)=181; N(Minority)=50; N(Caucasian)=208.

^a These questions were asked in 2008 and 2010 only, so the number of evaluated judges in each category is smaller: N(Women)=36; N(Men)=76; N(Minority)=23; N(Caucasian)=89.

FIGURE 7 :: ARIZONA :: Juror Evaluations of Trial Judges

	AVERAGE (0 – 4)					
	WOMEN	MEN	DIFFERENCE	MINORITY	CAUCASIAN	DIFFERENCE
INTEGRITY	3.7	3.7	0.0	3.7	3.7	0.0
Basic fairness and impartiality	3.7	3.7	0.0	3.8	3.7	0.0
Equal treatment regardless of race	3.7	3.7	0.0	3.7	3.7	0.0
Equal treatment regardless of gender	3.7	3.7	0.0	3.7	3.7	0.0
Equal treatment regardless of religion	3.7	3.7	0.0	3.7	3.7	0.0
Equal treatment regardless of national origin	3.7	3.7	0.0	3.7	3.7	0.0
Equal treatment regardless of disability	3.7	3.7	0.0	3.7	3.7	0.0
Equal treatment regardless of age	3.7	3.7	0.0	3.7	3.7	0.0
Equal treatment regardless of sexual orientation	3.7	3.7	0.0	3.7	3.7	0.0
Equal treatment regardless of economic status	3.7	3.7	0.0	3.7	3.7	0.0
COMMUNICATION SKILLS	3.7	3.7	0.0	3.8	3.7	0.0
Explained proceedings to the jury	3.8	3.8	0.0	3.8	3.8	0.0
Explained reasons for delays	3.6	3.7	0.0	3.7	3.7	0.0
Clearly explained the juror's responsibilities	3.8	3.8	0.0	3.8	3.8	0.0
JUDICIAL TEMPERAMENT	3.8	3.7	0.0	3.8	3.7	0.0
Understanding and compassion	3.7	3.7	0.0	3.7	3.7	0.0
Dignified	3.8	3.8	0.0	3.8	3.8	0.0
Courteous	3.8	3.8	0.0	3.8	3.8	0.0
Conduct that promotes public confidence in the court and judge's ability	3.8	3.8	0.0	3.8	3.8	0.0
Patient ^a	3.7	3.7	0.0	3.7	3.7	0.0
ADMINISTRATIVE PERFORMANCE	3.6	3.6	0.0	3.6	3.6	0.0
Punctual(ity) in conducting proceedings	3.5	3.4	0.0	3.4	3.5	0.0
Maintained (maintenance of) proper control of (over) courtroom	3.8	3.8	0.0	3.8	3.7	0.0
Was prepared for the proceedings ^a	3.7	3.8	-0.1	3.7	3.8	-0.1

N(Women)=41; N(Men)=119; N(Minority)=30; N(Caucasian)=130.

^aThese questions were asked in 2008 and 2010 only, so the number of evaluated judges in each category is smaller: N(Women)=21; N(Men)=47; N(Minority)=14; N(Caucasian)=54.

COLORADO

Colorado justices and judges have been evaluated since 1988. They are evaluated mid-term and prior to standing for retention. Here we analyze retention evaluation results from 2004 to 2010.

For supreme court justices and court of appeals judges, the State Commission on Judicial Performance surveys attorneys, court staff, other appellate judges, and district judges. For district court and county court judges, local commissions survey attorneys, non-attorneys (including jurors, witnesses, litigants, court staff, court interpreters, law enforcement personnel, crime victims, and social service case workers), and appellate judges.

Surveys include a variety of questions related to case management, application and knowledge of the law, communications, demeanor, and diligence. Respondents are asked to assign justices and judges a grade for each question of A (4), B (3), C (2), D (1), or F (0), and average grades are reported.

Because virtually identical questionnaires are used in evaluations of district court and county court judges, we combine their results in **Figures 8** (pg. 20) and **9** (pg. 21).⁵¹ The only difference in evaluations of judges of these courts is that an additional question is asked of attorneys regarding district court judges—i.e., “Providing written communications that are clear, thorough and well reasoned.”

With respect to attorney evaluations of trial judges, as seen in **Figure 8**, average scores for performance categories range from a low of 3.1—for women and minority judges on Application and Knowledge of the Law, to a high of 3.5—the average for all four groups on Communications. In terms of GPAs, average scores ranged from a B to a B+/A-.

Disparities in attorney evaluation scores for women and men trial judges are no greater than one-tenth of a point. The same is true for minority and Caucasian judges, with the exception of two questions for which minority judges score two-tenths of a point lower—“Willing to reconsider error in fact or law” and “Being willing to handle cases on the docket even when they are complicated and time consuming.”

Figure 9 shows that the range of “grades” from non-attorneys is slightly higher than from attorneys. Women, men, and Caucasian judges earn the highest GPA of 3.6 for Communications, while minorities receive the “lowest” GPA of 3.3—i.e., a B +—for Application of the Law. There are no differences based on gender in non-attorney evaluations of trial judges, and differences based on ethnicity are no greater than one-tenth of a point.

Figure 10 (pg. 21) displays appellate judges’ assessments of the “overall performance” of district court judges. Averages range from 3.4 for minority judges to 3.6 for women and Caucasian judges.

FIGURE 8 :: COLORADO :: Attorney Evaluations of Trial Judges

	AVERAGE (0 – 4)					
	WOMEN	MEN	DIFFERENCE	MINORITY	CAUCASIAN	DIFFERENCE
CASE MANAGEMENT	3.3	3.4	-0.1	3.3	3.4	-0.1
Promptly issuing a decision on the case after trial.	3.4	3.5	-0.1	3.4	3.5	0.0
Maintaining appropriate control over proceedings.	3.4	3.5	-0.1	3.4	3.4	0.0
Promptly ruling on pre-trial motions.	3.3	3.3	-0.1	3.2	3.3	-0.1
Setting reasonable schedules for cases.	3.3	3.4	-0.1	3.2	3.4	-0.1
APPLICATION AND KNOWLEDGE OF THE LAW	3.1	3.2	-0.1	3.1	3.2	-0.1
Being able to identify and analyze relevant facts.	3.2	3.3	-0.1	3.2	3.3	-0.1
Basing decisions on evidence and arguments.	3.1	3.2	-0.1	3.2	3.2	-0.1
Willing to reconsider error in fact or law.	2.9	3.0	-0.1	2.8	3.0	-0.2
Issuing consistent sentences when the circumstances are similar.	3.2	3.3	-0.1	3.2	3.3	-0.1
COMMUNICATIONS	3.5	3.5	0.0	3.5	3.5	0.0
Making sure all participants understand the proceedings.	3.6	3.6	0.0	3.5	3.6	0.0
Providing written communications that are clear, thorough and well reasoned. ^a	3.3	3.4	-0.1	3.3	3.3	0.0
DEMEANOR	3.3	3.4	-0.1	3.3	3.4	-0.1
Giving proceedings a sense of dignity.	3.4	3.5	-0.1	3.4	3.5	-0.1
Treating parties (participants) with respect.	3.4	3.5	0.0	3.4	3.5	-0.1
Conducting (his/her) courtroom in a neutral manner.	3.3	3.4	-0.1	3.3	3.3	-0.1
Consistently applying laws and rules.	3.2	3.4	-0.1	3.2	3.3	-0.1
DILIGENCE	3.3	3.3	-0.1	3.2	3.3	-0.1
Using good judgment in application of relevant law and rules.	3.2	3.2	-0.1	3.1	3.2	-0.1
Doing the necessary homework and being prepared for his/her cases.	3.3	3.4	-0.1	3.2	3.3	-0.1
Being willing to handle cases on the docket even when they are complicated and time consuming.	3.4	3.5	-0.1	3.3	3.4	-0.2

N(Women)=111; N(Men)=291; N(Minority)=49; N(Caucasian)=353.

^a This question is asked for district court judges only, so the number of evaluated judges is smaller: N(Women)=70; N(Men)=188; N(Minority)=28; N(Caucasian)=130.

FIGURE 9 :: COLORADO :: Non-Attorney Evaluations of Trial Judges

	AVERAGE (0 – 4)					
	WOMEN	MEN	DIFFERENCE	MINORITY	CAUCASIAN	DIFFERENCE
DEMEANOR	3.5	3.5	0.0	3.4	3.5	-0.1
Giving court proceedings a sense of dignity.	3.5	3.5	0.0	3.4	3.5	-0.1
Treating participants (in the case) politely and with respect.	3.5	3.5	0.0	3.4	3.5	-0.1
Conducting (his/her) court(room) in a neutral manner.	3.5	3.5	0.0	3.4	3.5	-0.1
Having a sense of compassion and human understanding for those who appear before the court.	3.4	3.5	0.0	3.4	3.5	-0.1
FAIRNESS	3.4	3.4	0.0	3.4	3.4	-0.1
Giving participants an opportunity to be heard.	3.4	3.4	0.0	3.4	3.4	-0.1
Treating those involved in the case without bias.	3.4	3.4	0.0	3.4	3.4	-0.1
Treating fairly people who represent themselves.	3.4	3.5	0.0	3.4	3.5	-0.1
Giving each side enough time to present his or her case.	3.4	3.4	0.0	3.4	3.4	-0.1
COMMUNICATIONS	3.6	3.6	0.0	3.5	3.6	-0.1
Making sure participants understand the proceedings, and what is going on in the courtroom.	3.5	3.5	0.0	3.4	3.5	-0.1
Using language that everyone can understand.	3.6	3.6	0.0	3.5	3.6	-0.1
Speaking clearly so everyone in the courtroom can hear what is being said.	3.6	3.6	0.0	3.5	3.6	-0.1
DILIGENCE	3.5	3.5	0.0	3.4	3.5	-0.1
Beginning court on time.	3.5	3.5	0.0	3.5	3.5	0.0
Maintaining appropriate control over proceedings.	3.5	3.5	0.0	3.4	3.5	-0.1
Setting reasonable schedules for cases.	3.4	3.4	0.0	3.4	3.4	-0.1
Being prepared for (his/her) cases.	3.5	3.5	0.0	3.5	3.5	0.0
Managing court proceedings so that there is little wasted time.	3.4	3.5	0.0	3.4	3.5	-0.1
APPLICATION OF LAW	3.4	3.4	0.0	3.3	3.4	-0.1
Giving reasons for rulings.	3.4	3.4	0.0	3.4	3.4	0.0
Willing to make decisions without regard to possible outside pressure.	3.4	3.4	0.0	3.3	3.4	-0.1
Being able to identify and analyze relevant facts.	3.4	3.4	0.0	3.3	3.4	-0.1

N(Women)=111; N(Men)=292; N(Minority)=49; N(Caucasian)=354.

FIGURE 10 :: COLORADO :: Appellate Judge Evaluations of District Judges

	AVERAGE (0 – 4)					
	WOMEN	MEN	DIFFERENCE	MINORITY	CAUCASIAN	DIFFERENCE
Overall performance as a judge.	3.6	3.5	0.1	3.4	3.6	-0.1

N(Women)=29; N(Men)=82; N(Minority)=11; N(Caucasian)=100.

UTAH

Utah's JPE program was established in 1986, with evaluations first issued in 1990. Appellate justices and judges and trial judges are evaluated mid-term and prior to standing for retention. Attorney evaluations of judges standing for retention from 2002 to 2010 are analyzed here.⁵²

For supreme court justices and court of appeals judges, the Judicial Performance Evaluation Commission surveys attorneys; for district court and juvenile court judges, surveys are sent to attorneys; for district court judges, surveys are also sent to jurors.⁵³ Attorney respondents are asked to rate justices and judges as Excellent, More Than Adequate, Adequate, Less Than Adequate, and Inadequate in several performance areas. Average percentages for each category are reported, as well as the percentage of Favorable responses, representing the sum for Excellent, More Than Adequate, and Adequate. Jurors are asked a series of Yes/No questions regarding whether judges exhibit a particular quality or behavior.

District court and juvenile court judges fare well in evaluations by attorneys, as shown in **Figure 11** (pgs. 23 and 24).⁵⁴ Evaluations of overall performance range from 90 percent Favorable for women to 93 percent Favorable for men. Favorable responses were at a high of 98 percent for male and Caucasian judges on "Maintains order in the courtroom." On the low end, for "Behavior is free from bias and favoritism," 87 percent of attorneys gave women judges a Favorable rating and 88 percent gave minority judges a Favorable rating.

The largest gap in ratings for women and men judges—five percent—exists for "Avoids ex parte communications," where men were rated Excellent slightly more often. Disparities in attorney ratings for minority and Caucasian judges are no greater than three percent.

“

BY FOCUSING ON NEUTRAL,
NONPARTISAN, PROCESS-ORIENTED
STANDARDS, RATHER THAN ON
PARTICULAR CASE OUTCOMES, JPE
ENSURES THAT JUDGES CAN FOLLOW
THE RULE OF LAW WHILE BEING HELD
TO BROADLY ACCEPTED STANDARDS FOR
THEIR OVERALL PERFORMANCE.

”

FIGURE 11 :: UTAH :: Attorney Evaluations of Trial Judges

		WOMEN AND MEN JUDGES					
		TOTAL FAVORABLE	EXCELLENT	MORE THAN ADEQUATE	ADEQUATE	LESS THAN ADEQUATE	INADEQUATE
Behavior is free from impropriety and the appearance of impropriety.	Women	92%	46%	30%	16%	5%	3%
	Men	94%	50%	28%	16%	4%	2%
	Difference	-2%	-4%	2%	1%	1%	1%
Behavior is free from bias and favoritism.	Women	87%	42%	28%	17%	8%	5%
	Men	91%	46%	29%	16%	6%	4%
	Difference	-4%	-4%	-1%	1%	3%	2%
Avoids ex parte communications.	Women	94%	45%	31%	18%	3%	2%
	Men	96%	50%	30%	17%	2%	1%
	Difference	-2%	-5%	1%	2%	1%	1%
Understands the rules of procedure and evidence.	Women	91%	40%	30%	22%	6%	3%
	Men	93%	41%	31%	21%	5%	2%
	Difference	-2%	-1%	-1%	1%	1%	1%
Is prepared for hearings and trials.	Women	95%	44%	32%	20%	3%	2%
	Men	95%	45%	31%	19%	3%	2%
	Difference	0%	-1%	0%	0%	0%	1%
Demonstrates appropriate demeanor.	Women	89%	47%	26%	17%	7%	5%
	Men	92%	49%	28%	15%	5%	3%
	Difference	-3%	-3%	-2%	1%	2%	2%
Maintains order in the courtroom.	Women	96%	46%	31%	20%	2%	2%
	Men	98%	50%	32%	16%	1%	1%
	Difference	-2%	-4%	-1%	4%	1%	1%
Issues orders and opinions without unnecessary delay.	Women	94%	39%	34%	21%	3%	3%
	Men	95%	40%	33%	23%	3%	2%
	Difference	-1%	-1%	1%	-1%	1%	1%
Effectively uses pretrial procedures to narrow and define the issues.	Women	94%	36%	33%	26%	4%	2%
	Men	94%	38%	33%	24%	4%	2%
	Difference	0%	-2%	0%	2%	0%	0%
Overall, the performance of this judge or commissioner is:	Women	90%	41%	30%	19%	6%	4%
	Men	93%	44%	32%	17%	4%	3%
	Difference	-3%	-3%	-2%	2%	2%	1%

MINORITY AND CAUCASIAN JUDGES							
		TOTAL FAVORABLE	EXCELLENT	MORE THAN ADEQUATE	ADEQUATE	LESS THAN ADEQUATE	INADEQUATE
Behavior is free from impropriety and the appearance of impropriety.	Minority	92%	50%	27%	16%	5%	3%
	Caucasian	94%	49%	29%	16%	4%	3%
	Difference	-2%	1%	-2%	0%	2%	0%
Behavior is free from bias and favoritism.	Minority	88%	46%	26%	16%	9%	4%
	Caucasian	91%	45%	29%	17%	6%	4%
	Difference	-3%	1%	-3%	-1%	3%	0%
Avoids ex parte communications.	Minority	95%	49%	29%	16%	4%	2%
	Caucasian	96%	49%	30%	17%	2%	2%
	Difference	-1%	0%	-1%	-1%	1%	0%
Understands the rules of procedure and evidence.	Minority	92%	41%	29%	22%	6%	2%
	Caucasian	93%	41%	31%	21%	5%	3%
	Difference	-1%	0%	-2%	1%	1%	0%
Is prepared for hearings and trials.	Minority	96%	46%	31%	18%	2%	2%
	Caucasian	95%	44%	31%	19%	3%	2%
	Difference	1%	2%	0%	-1%	-1%	0%
Demonstrates appropriate demeanor.	Minority	90%	50%	25%	15%	6%	4%
	Caucasian	92%	49%	28%	16%	5%	3%
	Difference	-2%	1%	-2%	-1%	1%	1%
Maintains order in the courtroom.	Minority	97%	48%	32%	17%	2%	1%
	Caucasian	98%	49%	32%	17%	1%	1%
	Difference	-1%	-1%	-1%	0%	1%	0%
Issues orders and opinions without unnecessary delay.	Minority	95%	42%	31%	22%	3%	2%
	Caucasian	95%	39%	33%	22%	3%	2%
	Difference	0%	2%	-2%	0%	0%	0%
Effectively uses pretrial procedures to narrow and define the issues.	Minority	94%	39%	31%	25%	4%	2%
	Caucasian	94%	37%	33%	24%	4%	2%
	Difference	0%	1%	-2%	1%	1%	0%
Overall, the performance of this judge or commissioner is:	Minority	91%	44%	29%	18%	5%	3%
	Caucasian	92%	43%	32%	17%	5%	3%
	Difference	-1%	1%	-3%	1%	1%	0%

N(Women)=28; N(Men)=129; N(Minority)=19; N(Caucasian)=138.

DISCUSSION

It is important to note at the outset that the results discussed in this report confirm the high quality of the justices and judges serving in these four states. While we find some differences in evaluation scores based on gender and ethnicity, these differences are slight—typically no more than one-tenth of a point or no larger than three percent, depending on the scale of measurement used. At the same time, these small disparities tend to be pervasive.

We cannot definitively attribute these score differences to one factor or another. Past research on differences between JPE scores for female/minority judges and male/Caucasian judges has considered a number of factors that may contribute to differences in assessments of judicial performance based on gender and ethnicity. Scholars have examined a variety of indicators of judicial potential and judicial quality, such as quality of legal education, extent of judicial experience, disciplinary history, and reversal rates. They have also speculated that there may be problems with the evaluation process itself, such as poorly constructed surveys, an inadequate pool of respondents, and respondent bias. Here we explore two of these possible factors further: extent of judicial experience and survey respondent bias.

EXTENT OF JUDICIAL EXPERIENCE

It is logical to expect that there may be a connection between a particular judge's performance evaluation scores and the length of her tenure on the bench—that performance may improve as a judge gains more experience. This may be particularly true in the four states examined here—i.e., states with JPE programs, where judges regularly receive constructive feedback to facilitate self-improvement. On the other hand, some have hypothesized that judges may experience a “burnout” after several years on the bench that affects the quality of their performance. In addition, evaluators may “take it easy” on newer judges while holding experienced judges to a higher standard.

In any event, the extent of a judge's experience on the bench may have an impact on his performance evaluation scores. To assess whether this is a factor for judges serving in the four states in this study, we calculated correlations between overall evaluation scores and the number of years served before standing for retention. In **Figure 12** (pg. 26), we report correlations for women, men, minority, Caucasian, and all judges for each category of respondents and level of court.

The correlation coefficient may range from -1.00 to +1.00. A negative correlation coefficient indicates that the overall evaluation score decreases the longer a judge has served on the bench, while a positive correlation coefficient means that a judge's overall evaluation score improves during her tenure on the bench.

While some of our correlation coefficients are negative, most are less than |0.20|, denoting a negligible relationship. There are, however, two exceptions. In Utah, there is a moderate negative correlation (-0.44) between attorney evaluations of minority trial judges and the extent of those judges' experience, suggesting that evaluation scores go down the longer a judge serves. In Alaska, there is a moderate negative correlation (-0.43) between juror evaluations of male judges and the length of those judges' tenure, indicating that evaluation scores decline over the course of a judge's tenure.

In terms of positive correlations, most represent only a negligible relationship between evaluation scores and length of judicial service, but stronger ties are seen in three areas. In Colorado, there is a moderate positive correlation (0.53) between appellate judge evaluations of minority trial judges and the number of years those judges have served, indicating that evaluation scores for these judges improve throughout their service. There is also a slight positive correlation (0.29)

between attorney evaluations of male judges in Alaska and the length of their tenure. Finally, in Arizona, there is a slight positive correlation (0.26) between attorney evaluations of minority trial judges and the extent of those judges' experience.

This analysis indicates that in some states and for some judges, the extent of judicial experience is a factor that has at least a small impact on judicial evaluation scores. As such, experience should be taken into account in interpreting performance evaluation results for these judges and in drawing comparisons between them and their colleagues on the bench. However, it seems clear that the length of a judge's tenure does not systematically influence the evaluation of his performance; rather, it is a minor factor that is relevant in a few discrete contexts.

FIGURE 12 :: Correlations: Overall Evaluation Scores and Years of Experience

	ATTORNEYS					NON-ATTORNEYS				
	WOMEN	MEN	MINORITY	CAUCASIAN	ALL	WOMEN	MEN	MINORITY	CAUCASIAN	ALL
Alaska^a										
All	-0.05 (13)	0.29 (58)	--	--	0.16 (71)	-0.01 (13)	0.21 (58)	--	--	0.17 (71)
Arizona										
Appellate	-0.11 (10)	-0.17 (31)	--	--	-0.15 (41)	n/a	n/a	n/a	n/a	n/a
Trial	0.10 (87)	0.13 (200)	0.26 (58)	0.09 (229)	0.13 (287)	-0.06 (77)	0.12 (181)	0.08 (50)	0.06 (208)	0.07 (258)
Colorado										
Trial	0.06 (111)	-0.01 (291)	-0.15 (49)	0.01 (353)	0.01 (402)	-0.03 (111)	-0.03 (292)	-0.01 (49)	-0.03 (354)	-0.02 (403)
Utah										
Trial	-0.18 (28)	-0.14 (129)	-0.44 (19)	-0.11 (138)	-0.14 (157)	n/a	n/a	n/a	n/a	n/a
	OTHER JUDGES					JURORS				
	WOMEN	MEN	MINORITY	CAUCASIAN	ALL	WOMEN	MEN	MINORITY	CAUCASIAN	ALL
Alaska										
Trial	n/a	n/a	n/a	n/a	n/a	0.11 (12)	-0.43 (55)	--	--	0.01 (67)
Arizona										
Appellate	-0.19 (10)	0.07 (31)	--	--	0.02 (41)	n/a	n/a	n/a	n/a	n/a
Trial	n/a	n/a	n/a	n/a	n/a	-0.07 (41)	-0.19 (119)	0.18 (30)	-0.20 (130)	-0.15 (160)
Colorado										
Trial	-0.04 (29)	0.04 (82)	0.53 (11)	-0.04 (100)	0.01 (111)	n/a	n/a	n/a	n/a	n/a
Utah										
Trial	n/a	n/a	n/a	n/a	n/a	--	--	--	--	--

Numbers in parentheses represent number of judges.

^a Non-attorneys include court employees only.

SURVEY RESPONDENT BIAS

Many prior studies in this area theorize that respondent biases are to blame for the differences in scores between female/minority judges and male/Caucasian judges:⁵⁵ Professor Sterling speculated that gender stereotypes were influencing disparities in Colorado JPE results in the early 1990s; Professor Burger's analysis of attorney evaluations of Missouri judges suggested that bias on the part of respondents was among the possible explanations for the discrepancies in scores; and the recent analysis of the *Las Vegas Review Journal* "Judging the Judges" bar poll found results to be "consistent with the hypothesis that judicial performance evaluation surveys may carry with them unexamined and unconscious gender and race biases."⁵⁶

If respondent bias *is* disproportionately disadvantaging female and minority judges, it is more likely to be implicit, rather than explicit, bias. An individual's unconscious biases begin developing at a very young age, and they do so automatically. According to Professor Jerry Kang, humans employ schemas, which are cognitive structures "that represent[] knowledge about a concept or type of stimulus, including its attributes and the relations among those attributes."⁵⁷ Further, we do so out of necessity.⁵⁸ This process is not inherently negative; humans classify people, things, and events in order to facilitate the processing of information. From this process, schema emerge that govern how we perceive gender and ethnicity. These stereotypes—associations "between a given object and a specific attribute"⁵⁹—emerge early in life (as young as three) and are learned from multiple sources, such as family, friends, community, media, etc.⁶⁰ Notably, stereotypes can involve favorable or unfavorable associations with certain character traits.⁶¹ For example, with respect to gender, implicit stereotypes present women as homemakers, caretakers, committed to family, etc., and with these traditional roles, stereotypical work styles and character traits are assigned—nurturing, caring, passive, sensitive, and gentle.⁶² These are not, in and of themselves, unfavorable character traits, but can become so when ascribed and applied wholesale to a group.

Once ingrained in adults, stereotypes affect how people perceive the world.⁶³ The tendency to invoke them in making judgments about people is spontaneous and automatic—they become implicit.⁶⁴ In recent decades, these implicit biases may have replaced some of the more overt forms of discrimination against women and minorities.⁶⁵ Of course, this does not mean that these implicit biases have any fewer real-world consequences than overt discrimination. Returning to the example of implicit gender stereotypes above, in the workplace these role and character stereotypes can affect how women are perceived in multiple ways. For example, a woman may be perceived as less committed to her career, instead favoring her family life, and her career and advancement opportunities may be impacted accordingly.⁶⁶ These stereotypes may also affect how a woman is perceived in the workplace when she fails to act in accordance with them—either because she naturally does not conform to the stereotypes attributed to her or because she is overcompensating for them, or both.⁶⁷

Numerous studies, using a variety of different methodologies, have "converged on the common finding that implicit gender stereotypes emerge in people's judgments and decisions when measured in ways that bypass decision makers' awareness of potential bias."⁶⁸ These biases affect one's judgment of others and can alter the way in which one remembers information.⁶⁹ Stereotypes tend to be more salient and influential when an occupation is culturally associated with a particular gender or ethnicity and where women or minorities are fairly rare, such as in leadership roles like

that of a judge.⁷⁰ In such roles, traditional stereotypes are supplemented and magnified by the stereotypical association between leadership roles and masculinity (with respect to gender) and leadership roles and Caucasians (with respect to ethnicity).⁷¹

The gender or ethnic identity of a profession effectively shapes the expected behavior of individuals in that profession.⁷² A repeated association between the Caucasian male and the profession of judge shapes expectations of competence.⁷³ If the job of judge is perceived to be stereotypically masculine, that perception would activate assumptions that associate competence with masculinity, so that men are perceived to be more competent than women.⁷⁴ Regarding ethnicity, if individuals create an association between Caucasian attributes (for example, assertive, ambitious, competitive), masculinity, and physical appearance, then as a consequence, Caucasian male qualities become the lodestar of a successful judge.⁷⁵ Women and minorities may experience a “backlash” when exercising authority in Caucasian male-dominated domains in a manner that is perceived as violating stereotype-based expectations.⁷⁶ For example, studies show that when female leaders behave in a directive, autocratic style, there is a tendency to evaluate them more negatively.⁷⁷ In this sense, there are conflicting expectations for female and male judges: “A male judge who strictly controls his courtroom runs a tight ship. His female counterpart is a bitch.”⁷⁸

Over the last few decades, countless committees and task forces on gender and racial bias around the country explored how gender and ethnicity affect the experiences and treatment of individuals who interact with the judicial system, including judges, attorneys, litigants, and court employees. Some of these committees considered gender bias only and left the question of racial and ethnic bias to a separate task force, while others considered the question of bias more generally by focusing on both gender and ethnicity. Although each committee grappled with its own definition of bias and how to identify it—while some made a point of not defining it at all, many definitions were based on stereotypical notions of women and minorities.⁷⁹ Using a combination of qualitative and quantitative information (techniques varied by committee), many of these committees found that bias was prevalent in the experiences of female and minority litigants, court employees, and attorneys.⁸⁰

At the same time, a number of committees documented gender and/or ethnic bias against judges. For example, in a 1989 publication of the Washington State Task Force on Gender and Justice in the Courts, the task force reported that 29 percent of attorney respondents to a survey indicated that affidavits of prejudice were used to disqualify female judges primarily because of their gender.⁸¹ A 1990 report of the Colorado Supreme Court’s Task Force on Gender Bias found evidence that female judges experienced hostility and a lack of respect both from counsel before them and from their colleagues.⁸² A 1997 report of the Second Circuit Task Force on Gender, Racial and Ethnic Fairness in the Courts found that seven percent of female judges reported hearing disparaging remarks about female judges, and 11 percent reported similar remarks about minority judges.⁸³

Evidence of implicit biases against women and minorities in the legal profession has since been studied in more detail in the social sciences. For example, using various components, Justin Levinson and Danielle Young undertook a study of law students at the University of Hawai’i to explore whether the students viewed a judge as an “implicit male leader prototype.”⁸⁴ The pair found that law students held implicit gender biases related to judges, showing a significant association between judge and male, compared to that found between judge and female.⁸⁵ Similar findings have emerged

from other studies. Lynn Hecht Schafran has called the Caucasian male “central casting” for a judge: “The status set for judges is still white and male. Thus, white women judges are one step removed from the ‘norm.’ Women of color judges are two steps removed.”⁸⁶

With respect to our study, some of the areas in which we found greater disparities between scores for women and men judges, and minority and Caucasian judges, were areas in which past research suggests that judges may be held to different standards based on gender and/or ethnicity. For example, Alaska attorneys rated women justices and judges lower than their male colleagues on their impartiality/fairness and their temperament. In Arizona, minority trial judges received lower scores from attorneys on both their legal ability and their communication skills, and from non-attorneys on their judicial temperament. These areas in performance evaluation surveys invoke respondents’ expectations with respect to appropriate judicial demeanor and their perceptions of judicial competence. It is possible, then, that some of the gender- and ethnicity-based differences we found in judicial performance evaluation scores are a product of implicit biases.

“

[S]OME OF THE AREAS IN WHICH WE FOUND GREATER DISPARITIES BETWEEN SCORES FOR WOMEN AND MEN JUDGES, AND MINORITY AND CAUCASIAN JUDGES, WERE AREAS IN WHICH PAST RESEARCH SUGGESTS THAT JUDGES MAY BE HELD TO DIFFERENT STANDARDS BASED ON GENDER AND/OR ETHNICITY.

”

RECOMMENDATIONS

Because implicit bias is a serious issue that has the potential to impact JPE—and to do so in ways that are difficult to detect—both existing and developing JPE programs should strive to minimize the potential impact of such bias to the extent possible. This can be done in a number of ways: 1) assure that the data collection effort is broad and deep, with a variety of data sources, and that the committee that synthesizes the data is itself diverse; 2) ensure that JPE survey instruments represent the best thinking in survey design informed by cognitive psychology, in order to reduce opportunities for respondents’ implicit biases to come through; and 3) educate the bench, bar, court staff, and the public broadly on implicit bias and ways in which to address and minimize it.

COMPREHENSIVE JPE PROGRAMS

IAALS recommends a comprehensive JPE program as one means of accounting for, and mitigating, potential implicit biases—specifically, one that takes into account a variety of data sources and involves a diverse commission in vetting the results.⁸⁷ According to Kevin Esterling, evaluations by performance evaluation commissions “differ from the more familiar bar polls and interest group judicial ratings, and they do so in ways that are likely to make the performance information more credible or persuasive to both judges and citizens.”⁸⁸ The world of potential data sources that can be incorporated into JPE programs is considerable, and the hallmark of comprehensive JPE programs is that they employ several of the following evaluation tools: attorney surveys, non-attorney surveys, court staff surveys, trial and/or appellate judge surveys, peer judge surveys, self-evaluation, courtroom observation, recusal and reversal rates, disciplinary records, written opinion review, public comment, interviews with the judge being evaluated, and caseload management statistics. In contrast, unofficial JPE programs and standalone bar polls typically utilize surveys of attorneys only. Such programs “lack the key trappings of comprehensive JPE programs” and “[b]ar surveys alone cannot account properly, if at all, for the reaction of jurors, litigants, witnesses, and court staff; the judge’s ability as a case manager; or impartial courtroom observation.”⁸⁹

The breadth of data available for consideration in the more comprehensive JPE programs can mitigate the impact of potential implicit biases that could be invoked in JPE surveys by providing additional data on which to base performance evaluations. In late 2010, the Colorado State Commission on Judicial Performance conducted an assessment of state and local judicial performance evaluation processes, in order to gather JPE commissioners’ perspectives on the evaluation process.⁹⁰ Respondent commissioners were asked what information is used in deciding upon retention recommendations, the extent to which they consider information from JPE surveys, and what information from the surveys is most valuable. Of all the information considered by commissioners,⁹¹ the two sources considered strongly influential by over two-thirds of respondents were the judge interview (81 percent) and courtroom observation (68 percent); just over half of respondents considered survey results strongly influential. Regarding survey data, smaller majorities of respondents reported that survey comments (56 percent), overall survey results (52 percent), and the number of survey responses received (51 percent) are also influential. With respect to the importance of various data points in surveys, according to commissioners, the percentage of attorneys recommending retention, attorney comments, and attorneys’ average grades for judges were “strongly influential” less often than other factors (48 percent, 42 percent, and 42 percent of the time; respectively). Notably, only 44 percent of respondents indicated that numerical data from surveys was strongly influential.

These findings are important in a number of respects. First, they demonstrate that JPE commissioners rely upon sources of data apart from survey results in making recommendations regarding retention—and that commissioners find that data to be influential more often than the numerical data from surveys. Second, while a majority of commissioner respondents indicated that survey comments are strongly influential, less than a majority found *attorney* comments to be strongly influential. In fact, data from attorney surveys—retention recommendation, comments, and average grade given—did not score as highly as other sources of information. Because the greatest discrepancies in JPE data between female/minority scores and male/Caucasian scores are generally found in attorney evaluations, the fact that commissioners do not find data from attorney surveys to be particularly influential, at least with respect to other sources, suggests that potential implicit biases might be mitigated in JPE programs that incorporate the views of multiple respondent groups.

WELL-DESIGNED JPE SURVEYS

The quality of the survey instrument shapes the quality of the data collected. Different forms of biases find their way into survey results through a variety of access points—for example, from procedures used in identifying potential respondents and disseminating surveys, from the survey instrument itself (structure, design, question wording, scale used, etc.), and from individual responses.⁹² Given these potential sources of bias, it is imperative that JPE programs utilize the best practices in survey design when developing survey instruments.

In early 2010, the Administrative Office of the Illinois Courts moved toward a new JPE program. The prior JPE program for circuit and associate judges was voluntary and confidential, intended to “promote judicial excellence and competence.”⁹³ Under the new program, all circuit and associate judges are required to undergo performance evaluation at least once during their time on the bench. The National Center for State Courts (NCSC) serves as the administrator of the Illinois program. As part of this role, the NCSC developed new survey instruments and subjected them to rigorous testing to ensure their reliability and validity. This process consisted of an initial review of JPE surveys currently used by states across the country, as well as of national JPE guidelines and recommendations.

Based on this review, the NCSC developed a new instrument, focusing on item and response-scale clarity and emphasizing concrete, observable behaviors. The NCSC then solicited feedback through the following processes: a review by the Illinois Supreme Court Judicial Performance Evaluation Committee; consultation with academic experts on performance evaluation and design; a pilot test; and follow-up cognitive interviews with attorneys to explore their perspectives on the survey instrument. This feedback informed additional revisions to the survey.

A number of efforts were made during this process to ensure survey validity and reduce the potential impact of implicit bias.⁹⁴ Consultations with academic experts on bias in survey-based job performance evaluation suggested that additional behavior-based language be incorporated into the surveys and recommended that survey items encourage respondents to recall instances of observed behavior, rather than relying on assumptions, stereotypes, or other heuristics in formulating their answers. A frequency-based response scale was adopted in lieu of an agree/disagree response scale, the latter being more susceptible to common forms of response bias. The academic experts also recommended that the NCSC incorporate a structured free-recall (SFR) component into the surveys, calling on respondents to complete a memory-based task before beginning the actual survey—a proven technique in minimizing potential respondent biases. The SFR component of the survey 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.

Once fully vetted, revised, and finalized, the NCSC used the attorney survey to develop a survey for court staff, eliminating items related to behaviors that court staff would not have the opportunity to observe. Following the first year of the program, during which one-third of the relevant pool of judges were evaluated, the NCSC analyzed the results and “found no evidence of gender bias in the average performance ratings (total or by performance area) for male and female judges.”⁹⁵ A pilot study during this review was used to gauge the extent to which the SFR exercise reduced observed gaps between ratings of white male judges and minority female judges.⁹⁶ The survey containing the SFR component produced smaller differences between ratings than did a survey that did not incorporate the SFR questions, though these differences were not statistically significant.

The attorney and court staff survey instruments developed through this process are attached as Appendix B and Appendix C. Along with the NCSC, IAALS recommends these surveys, and the process through which they were developed, as models that may be tailored to fit the specific needs and objectives of particular states—whether to improve existing JPE programs or to develop new ones.

BROAD EDUCATION ON IMPLICIT BIAS

Because implicit biases based on deeply held stereotypes form outside an individual’s conscious awareness, they are especially difficult to detect and eliminate.⁹⁷ Indeed, implicit biases may operate in conjunction with one’s conscious efforts to be impartial and unbiased. There are, however, strategies that individuals and court systems can employ to address and reduce the influence of implicit bias. In recent years, the NCSC, through its National Campaign to Ensure the Racial and Ethnic Fairness of America’s State Courts, has devoted substantial time and resources to exploring the issue of implicit bias in the courts and developing educational materials on the issue.⁹⁸

With respect to reducing the influence of implicit bias, the NCSC has identified the following empirically-based strategies:⁹⁹

- Raise awareness of implicit bias.¹⁰⁰
- Seek to identify and consciously acknowledge real group and individual differences.
- Routinely check thought processes and decisions for possible bias.
- Identify distractions and sources of stress in the decision-making environment and remove or reduce them.
- Identify sources of ambiguity in the decision-making context and establish more concrete standards before engaging in the decision-making process.
- Institute feedback mechanisms.
- Increase exposure to stigmatized group members and counter-stereotypes and reduce exposure to stereotypes.

In the NCSC's *Addressing Implicit Bias in the Courts*, each strategy is explained in detail and accompanied by potential actions that both the individual and court can take in furtherance of the particular strategy.¹⁰¹ Although not all of the above strategies are directly relevant to the JPE context, the suggestions deserve broad consideration by the bench, bar, and individuals and organizations charged with evaluating judicial performance.

NCSC research has also identified conditions and risk factors that increase the likelihood that implicit bias might influence thoughts and decisions.¹⁰² These risk factors are as follows:

- Certain emotional states (for example, anger and disgust).
- Ambiguity—when the basis for judgment is somewhat vague.
- Salient social categories.
- Low-effort cognitive processing.
- Distracted or pressured decision-making circumstances.
- Lack of feedback that holds decision makers accountable for their judgments and actions.

Where these risk factors are present, the strategies listed above can also be helpful in mitigating the effects of implicit bias. Simply understanding and recognizing these risk factors can be an effective strategy in itself.

CONCLUSION

Administered properly, judicial performance evaluation is an invaluable means of fostering professional development for judges, informing voters and others responsible for deciding whether to retain judges in office, ensuring appropriate judicial accountability, and promoting and preserving public confidence in the courts. However, continuous review and improvement must be incorporated into a robust program. In order to maximize the benefits of JPE, IAALS suggests that JPE commissioners and program coordinators revisit their programs with a view toward utilizing multiple and diverse evaluation tools, developed in accordance with best practices in survey design and administration; involving a diverse group of individuals in overseeing the program; and taking into account the potential for implicit bias to impact evaluations. IAALS thanks and commends the states included in this study, and the states whose programs are in early stages or under development, for making this commitment to continuous program improvement and encourages jurisdictions considering an official JPE program to follow their lead. High-quality JPE programs promote high-quality judges—and that is the ultimate objective.

ENDNOTES

- ¹ INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., *THE BENCH SPEAKS ON JUDICIAL PERFORMANCE EVALUATION: A SURVEY OF COLORADO JUDGES* 13 (2008) [hereinafter *THE BENCH SPEAKS*].
- ² *Id.* at 14.
- ³ *Id.* at 13.
- ⁴ *Id.* at 28.
- ⁵ *Id.*
- ⁶ See David C. Brody, *The Use of Judicial Performance Evaluation to Enhance Judicial Accountability, Judicial Independence, and Public Trust*, 86 DENV. U. L. REV. 115, 127 (2008).
- ⁷ INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS. & LEAGUE OF WOMEN VOTERS OF COLO., *COLORADO VOTER OPINIONS ON THE JUDICIARY* 3 (2007).
- ⁸ Brody, *supra* note 6, at 127; INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., *SHARED EXPECTATIONS: JUDICIAL ACCOUNTABILITY IN CONTEXT* 13-14 (2006) [hereinafter *SHARED EXPECTATIONS*].
- ⁹ Jordan M. Singer, *The Mind of the Judicial Voter*, 2011 MICH. ST. L. REV. (forthcoming 2012) (SSRN version at 44).
- ¹⁰ *THE BENCH SPEAKS*, *supra* note 1, at 31-32.
- ¹¹ *Id.*
- ¹² Rebecca Love Kourlis & Jordan M. Singer, *Using Judicial Performance Evaluations to Promote Judicial Accountability*, 90 JUDICATURE 200, 206 (2007).
- ¹³ See generally INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., *QUALITY JUDGES INITIATIVE—IMPLEMENTATION*, <http://iaals.du.edu/initiatives/quality-judges-initiative/implementation/judicial-performance-evaluation> (providing detailed information on judicial performance evaluation in the states) (last visited June 7, 2012).
- ¹⁴ Press Release, N.C. Bar Ass'n, *NCBA Releases Judicial Performance Evaluation Report* (Jan. 3, 2012), <http://www.ncbar.org/about/communications/news/2012-news-articles/ncba-releases-judicial-performance-evaluation-report.aspx> (last visited June 7, 2012).
- ¹⁵ LAS VEGAS REVIEW-JOURNAL, *JUDGING THE JUDGES 2011 REPORT*, http://www.lvrj.com/hottopics/in_depth/judges (last visited June 7, 2012).
- ¹⁶ CIARA TORRES-SPELLISCY ET AL., *BRENNAN CTR. FOR JUSTICE, IMPROVING JUDICIAL DIVERSITY* 1 (2010) (footnotes omitted) (emphasis omitted).
- ¹⁷ *Id.* at 4.
- ¹⁸ Responses were received from 10,857 attorneys in the state of Colorado who were selected as part of a sample to evaluate judges before whom they had appeared during the previous 18 months. Joyce S. Sterling, *The Impact of Gender Bias on Judging: Survey of Attitudes Toward Women Judges*, 22 COLO. LAWYER 257, 257 (1993).
- ¹⁹ *Id.*
- ²⁰ *Id.*
- ²¹ *Id.*
- ²² Female attorney respondents ranked women judges as high as male judges on the items of “Displays adequate knowledge of substantive law” and “Displays adequate knowledge of the rules of evidence and procedure.” *Id.*
- ²³ *Id.*
- ²⁴ AM. BAR ASS'N, *COMM'N ON WOMEN IN THE PROFESSION—ABOUT THE COMMISSION*, http://www.americanbar.org/groups/women/about_us.html (last visited June 7, 2012).
- ²⁵ AM. BAR ASS'N, *COMM'N ON WOMEN IN THE PROFESSION, UNFINISHED BUSINESS: OVERCOMING THE SISYPHUS FACTOR* 14 (Dec. 1995).
- ²⁶ *Id.* at 14, 17.
- ²⁷ *Id.* at 17.
- ²⁸ Christine M. Durham, *Gender and Professional Identity: Unexplored Issues in Judicial Performance Evaluation*, 39 JUDGES J. 11, 12 (Spring 2000).

²⁹ *Id.* at 14.

³⁰ *Id.*

³¹ The commission formally rejected one judge (May 2004) and the other four judges retired after being informed they would not be reappointed (Nov. 2001, May 2002, Oct. 2002, Apr. 2003 and May 2003). Rob Perez, *Statistics spur some to say bench is gender biased*, HONOLULU STAR-BULLETIN, May 31, 2004.

³² *Id.*

³³ Ken Kobayashi, *State bar to share evaluations*, STAR ADVERTISER, Jan. 18, 2011.

³⁴ MOUND CITY BAR ASS'N, COMMUNITY AFFAIRS COMM., REPORT OF COMMUNITY AFFAIRS COMMITTEE REGARDING THE VOTERS' INFORMATION ABOUT JUDGES 5 (2007) (on file with author) [hereinafter REPORT OF COMMUNITY AFFAIRS COMM.]. An executive summary of the report is available in the June 2007 edition of the *Mound City News*.

³⁵ GARY K. BURGER, ATTORNEYS' RATINGS OF JUDGES: ST. LOUIS 1998-2006 5 (Feb. 15, 2007).

³⁶ These courts include the Missouri Court of Appeals, the 21st Judicial Circuit-St. Louis County, and the 22nd Judicial Circuit -St. Louis City. *Id.*

³⁷ *Id.*

³⁸ REPORT OF COMMUNITY AFFAIRS COMM., *supra* note 34, at 10.

³⁹ Professor Burger noted that, although concerning, the data “taken alone, do not argue in favor of one or another” explanation and that further information was needed to make definitive findings. ATTORNEY'S RATINGS OF JUDGES, *supra* note 35, at 6. In the wake of this study, at the recommendation of the Missouri Judicial Evaluation Survey Committee and with support from then-Chief Justice Michael Wolff, Missouri adopted an official JPE program by Supreme Court rule in 2008.

⁴⁰ Rebecca D. Gill et al., *Are Judicial Performance Evaluations Fair to Women and Minorities? A Cautionary Tale from Clark County, Nevada*, 45 LAW & SOC'Y REV. 731, 749 (2011).

⁴¹ *Id.* at 742-49, 751. The authors “concede that there are alternative objective measures of judge quality not included” in their model. *Id.* at 753.

⁴² *Id.* Of the 95 judges included in their study, 33 were women and five were racial or ethnic minorities. *Id.* at 741.

⁴³ *Id.* at 753, 755.

⁴⁴ Kansas, Missouri, New Mexico, and Tennessee also have official retention evaluation programs, but they are not included in this analysis. Data from only two retention cycles—2008 and 2010—were available for Kansas and Missouri. For New Mexico, only one retention cycle of data—2010—was available to us, including only four appellate judges and sixteen metropolitan court judges, and we could not ascertain the race/ethnicity of all judges standing for retention in that cycle. In Tennessee, the Judicial Performance Evaluation Commission only makes public a recommendation as to whether each appellate judge standing for retention should be retained or replaced, accompanied by a narrative—rather than numerical—summary of survey results.

⁴⁵ One of the strengths of each of these programs is that they strive continually to improve upon the process, in the interests of accurately capturing judges' performance on the bench and of providing the most useful information to voters. As such, survey questions may be revised from one retention cycle to the next, or more wholesale changes may be made.

⁴⁶ As our data consists of the population of judges who were evaluated in each cycle, rather than a random sample of judges, we do not conduct statistical hypothesis tests.

⁴⁷ Only one minority judge stood for retention from 2006 to 2010.

⁴⁸ Superior court judges in all other counties are chosen in contested elections and are not included in the performance evaluation program.

⁴⁹ Only four minority judges stood for retention from 2002 to 2010.

⁵⁰ KEVIN M. ESTERLING & KATHLEEN M. SAMPSON, JUDICIAL RETENTION EVALUATION PROGRAMS IN FOUR STATES: A REPORT WITH RECOMMENDATIONS—EXECUTIVE SUMMARY 8 (1998); *see generally* E. Lee Bernick & David J. Pratto, *A Behavior-based Evaluation Instrument for Judges*, 18 JUST. SYS. J. 173, 175 (1995) (detailing the “Halo Effect”).

⁵¹ For appellate courts, only five women judges and three minority judges stood for retention from 2004 to 2010.

⁵² In 2008, administration of Utah's JPE program transitioned from the Judicial Council to the Judicial Performance Evaluation Commission. With the transition, a number of changes were made to the program, including changes in survey content, expansion of survey respondent groups, and the introduction of a courtroom observation program. See Joanne C. Slotnik, *Warning: Your Name Might Be Listed Here*, 22 UTAH BAR J. 26 (2009).

⁵³ For justices and judges standing for retention in 2012, court staff were added to the respondent pool.

⁵⁴ Regarding appellate judges, only six women judges and no minority judges stood for retention from 2002 to 2010. Regarding juror evaluations of district court judges, survey questions for jurors changed fairly substantially in 2008, so only the 2008 and 2010 cycles were feasible for analysis. Only six women judges and seven minority judges stood for retention in 2008 and 2010.

⁵⁵ See *infra* pp. 6-8.

⁵⁶ Gill et al., *supra* note 40, at 749.

⁵⁷ Jerry Kang, *Trojan Horses of Race*, 118 HARV. L. REV. 1489, 1498 (2005).

⁵⁸ *Id.* at 1499.

⁵⁹ JERRY KANG, NAT'L CTR. FOR STATE COURTS, *IMPLICIT BIAS: A PRIMER FOR COURTS* 8 (August 2009).

⁶⁰ Justin D. Levinson, *Forgotten Racial Equality: Implicit Bias, Decisionmaking, and Misremembering*, 57 DUKE L.J. 345, 363 (2007).

⁶¹ Anthony G. Greenwald & Linda Hamilton Krieger, *Implicit Bias: Scientific Foundations*, 94 CAL. L. REV. 945, 951 (2006).

⁶² See Justin D. Levinson & Danielle Young, *Implicit Gender Bias in the Legal Profession: An Empirical Study*, 18 DUKE J. GENDER L. & POL'Y 1, 3 (2010); William T. Bielby, *Can I Get a Witness? Challenges of Using Expert Testimony on Cognitive Bias in Employment Discrimination Litigation*, 7 EMPL. RTS. & EMPLOY. POL'Y J. 377, 379 (2003).

⁶³ Levinson & Young, *supra* note 62, at 6.

⁶⁴ Bielby, *supra* note 62, at 379 (2003).

⁶⁵ See Cynthia Grant Bowman, *Women and the Legal Profession*, 7 AM. U. J. GENDER SOC. POL'Y & L. 149, 166 (1999) (noting that the 35 state and federal judicial circuits that had issued gender bias reports as of July 1998 concluded that "subtle forms of discrimination against women by both judges and male lawyers have largely, but not entirely, replaced more overt discrimination against women in litigation.").

⁶⁶ See Levinson & Young, *supra* note 62, at 9-12.

⁶⁷ *Id.* at 10-11; *Women on the Bench*, 12 COLUM. J. GENDER & L. 361, 362 (2003), panel held at Columbia Law School on October 19, 2002.

⁶⁸ David L. Faigman et al., *A Matter of Fit: The Law of Discrimination and the Science of Implicit Bias*, 59 HASTINGS L.J. 1389, 1410 (2008).

⁶⁹ See Levinson & Young, *supra* note 62, at 7-8.

⁷⁰ *Id.* at 12-13.

⁷¹ See *id.* at 13.

⁷² See Amy L. Wax, *Discrimination as Accident*, 74 IND. L.J. 1129, 1136-38 (1999).

⁷³ See Levinson & Young, *supra* note 62, at 12-13.

⁷⁴ See Faigman, *supra* note 68, at 1413-19.

⁷⁵ See *id.* at 1414; Ashleigh Shelby Rosette et al., *The White Standard: Racial Bias in Leader Categorization*, 93 J. OF APPLIED PSYCHOL. 758, 760 (2008).

⁷⁶ Cecilia L. Ridgeway, *Gender, Status, and Leadership*, 57 J. SOC. ISSUES 637, 649 (2001); Faigman, *supra* note 68, at 1413-15.

⁷⁷ See Ridgeway, *supra* note 76, at 648-49; see Faigman, *supra* note 68, at 1419-21.

⁷⁸ Lynn Hecht Schafran, *Not From Central Casting: The Amazing Rise of Women in the American Judiciary*, 36 U. TOL. L. REV. 953, 960 (2005).

⁷⁹ See, e.g., Jeannette F. Swent, *Gender Bias at the Heart of Justice: An Empirical Study of State Task Forces*, 6 S. CAL. REV. L. & WOMEN'S STUD. 1, 35-39 (1996) (comparing the various task forces' attempts to define gender bias).

⁸⁰ There were numerous, recurring reported examples of gender bias in the court system against women: women received unfavorable outcomes in cases due to gender, women's complaints were trivialized by the court more often than men's, women were the victims of demeaning behavior in the court, and women were treated as less credible. In general, white men believed that gender and race played little role in the court system. *Id.* at 55-70. See Bowman, *supra* note 65, at 165-68 (noting "that thousands of lawyers in many different states have reported virtually the same things.").

⁸¹ Schafran, *supra* note 78, at 959.

⁸² Sterling, *supra* note 18, at 257.

⁸³ Second Circuit Judicial Council Task Force on Gender, Racial and Ethnic Fairness, *Report of the Second Circuit Task Force*, 1997 ANN. SURV. AM. L. 117, 169 (1997).

⁸⁴ Levinson & Young, *supra* note 62, at 23, 28. Levinson and Young used two Implicit Association Tests (IATs) in their study. The Judge/Gender IAT, specifically, allowed Levinson and Young to test whether law students associate judges with males and paralegals with females.

⁸⁵ *Id.* at 28.

⁸⁶ Schafran, *supra* note 78, at 957.

⁸⁷ See generally INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., A BLUEPRINT FOR JUDICIAL PERFORMANCE EVALUATION (2006) (recommending 15 "building blocks" for a comprehensive judicial performance evaluation program).

⁸⁸ Kevin M. Esterling, *Judicial Accountability the Right Way*, 82 JUDICATURE 206, 209 (1999).

⁸⁹ SHARED EXPECTATIONS, *supra* note 8, at 61.

⁹⁰ OMNI INST., AN ASSESSMENT OF THE COLORADO JUDICIAL PERFORMANCE EVALUATION PROCESS (July 26, 2011) (on file with the Colorado State Commission on Judicial Performance).

⁹¹ Commissioners are required to consider the following sources of information: interview with judge, courtroom observation, survey results, judge statistics, review of judge decisions, and judge's self-evaluation. Discretionary sources of information include the following: judge's previous evaluation, public hearings, and comments from other interested individuals.

⁹² Bernick & Pratto, *supra* note 50, at 174.

⁹³ Ill. Sup. Ct. R. 58(b) (2012).

⁹⁴ See NAT'L CTR. FOR STATE COURTS, OVERVIEW OF THE SURVEY DEVELOPMENT PROCESS FOR THE SUPREME COURT OF ILLINOIS' JUDICIAL PERFORMANCE EVALUATION PROGRAM (2011) (on file with the National Center for State Courts).

⁹⁵ *Id.* at 2.

⁹⁶ JENNIFER ELEK & DAVID ROTTMAN, NAT'L CTR. FOR STATE COURTS, REPORT ON THE FIRST ROUND OF JUDICIAL PERFORMANCE EVALUATIONS 56-57 (June 22, 2011) (on file with the National Center for State Courts).

⁹⁷ KANG, *supra* note 59, at 1-2.

⁹⁸ NAT'L CTR. FOR STATE COURTS, RACE & ETHNIC FAIRNESS IN THE COURTS, http://www.ncsconline.org/D_Research/ref/implicit.html (offering NCSC project reports and resources on implicit bias).

⁹⁹ NAT'L CTR. FOR STATE COURTS, ADDRESSING IMPLICIT BIAS IN THE COURTS (2012) (summarizing the NCSC's project on implicit bias and judicial education, available at www.ncsc.org/ibreport).

¹⁰⁰ See, e.g., CAL. ADMIN. OFFICE OF THE COURTS, AO-TV SHOWS—THE NEUROSCIENCE AND PSYCHOLOGY OF DECISIONMAKING: A NEW WAY OF LEARNING, <http://www2.courtinfo.ca.gov/cjer/aotv/dialogue/neuro/index.htm> (last visited June 7, 2012) (offering a cost-free program that discusses "how unconscious processes may affect our decisions.").

¹⁰¹ NAT'L CTR. FOR STATE COURTS, STRATEGIES TO REDUCE THE INFLUENCE OF IMPLICIT BIAS (2012).

¹⁰² *Id.*

APPENDIX A :: State-by-State Overview of JPE Data

	YEARS	JUDGES EVALUATED	RESPONDENTS	CRITERIA	MEASUREMENT	SCORE
Alaska	2006-2010	Supreme Appeals	AK Bar Association members (attorneys and judges) Court employees	Legal ability (<i>bar association only</i>) Impartiality & fairness Integrity Temperament Diligence Overall evaluation	Poor (1) Deficient (2) Acceptable (3) Good (4) Excellent (5)	Average (1-5)
		Superior District	AK Bar Association members (<i>attorneys and judges</i>) Police/probation officers, social workers, court employees	Legal ability (<i>bar association only</i>) Impartiality & fairness Integrity Temperament Diligence Overall evaluation		
		Jurors		Fair and impartial to all sides Respectful and courteous to parties Attentive during proceedings Exercised control over proceedings Intelligence and skill as a judge Overall performance		
Arizona	2002-2010	Supreme Appeals	Attorneys Staff attorneys Superior court judges	Legal ability Integrity Communication skills at oral argument (<i>attorneys, staff attorneys only</i>) Judicial temperament (<i>attorneys, staff attorneys only</i>) Administrative performance	Superior (4) Very good (3) Satisfactory (2) Poor (1) Unsatisfactory (0)	Average (0-4)
		Superior	Attorneys Non-attorneys Jurors	Legal ability (<i>attorneys only</i>) Integrity Communication skills Judicial temperament Administrative performance Settlement skills (<i>attorneys only</i>)		
Colorado	2004-2010	Supreme Appeals	Attorneys District judges	<i>Series of statements & retention recommendation—no categories/criteria</i>	A (4) B (3) C (2) D (1) F (0)	Average (0-4)
		District County	Attorneys	Case management Application and knowledge of the law Communications Demeanor Diligence Retention recommendation		
			Non-attorneys (<i>including jurors</i>)	Demeanor Fairness Communications Diligence Application of law Retention recommendation		
			Appellate judges (<i>district judges only</i>)	<i>Overall grade</i>		
Utah	2002-2010	Supreme Appeals	Attorneys	Series of statements — no categories/criteria	Excellent More than adequate Adequate Less than adequate Inadequate	% Favorable
		District	Attorneys	Series of statements — no categories/criteria		
			Jurors (2008-2010)	Series of statements — no categories/criteria	Yes/no	% Yes
		Juvenile	Attorneys	Series of statements — no categories/criteria	Excellent More than adequate Adequate Less than adequate Inadequate	% Favorable

SUPREME COURT OF ILLINOIS JUDICIAL PERFORMANCE EVALUATION PROGRAM EVALUATION FOR ATTORNEY

You have been nominated by **Judge X** to participate in his/her judicial performance evaluation. This evaluation form contains a series of multiple choice and open-ended questions about **Judge X's** behavior along 5 dimensions of judicial performance that should take no more than 20 minutes to complete.

If you have any questions about this evaluation form, please contact our Help Desk specialist, XXXX XXXX, at XXX@XXX.org or 303-XXX-XXXX.

For more information about the Illinois Judicial Performance Evaluation Program, please contact the Illinois JPE program manager, XX XXXXXXXX XXXX, at XXXX@court.state.il.us or 217-XXX-XXXX.

YOUR APPEARANCES BEFORE JUDGE X

The following questions will help us understand the nature and extent of your experience with **Judge X**. Your responses are confidential and will not be used in any way that could identify you as a respondent.

1. Please estimate the number of your cases that **Judge X** has presided over in the past 12 months: _____ cases
2. Please estimate the number of times you appeared in court before **Judge X** in the past 12 months: _____ appearances*

***If you answered "0" to Question 2, STOP HERE** and return this evaluation form to XXXX XXX by email (XXX@XXXX.org) or fax (303-XXX-XXX, ATTN: XXXX XXX).

Only those attorneys who have appeared before the evaluated judge within the past 12 months should continue to page 2 of this evaluation form.

3. Regarding your appearances before **Judge X**, what was the nature of the **majority** of those cases?

- Civil
- Criminal - Adult
- Criminal - Juvenile
- Other: _____

4. Regarding your appearances before **Judge X**, which party did you **most often** represent?

- Civil Plaintiff
- Civil Defendant
- Criminal Prosecution
- Criminal Defense
- Equal Mix

5. Overall, of your cases before **Judge X** that went to trial and decision, **Judge X** ruled...

- Primarily in your party's favor
- About equally in favor of and against your party
- Primarily against your party
- N/A

THE DIMENSIONS OF JUDICIAL PERFORMANCE

In the subsequent sections of this evaluation form, you will be asked to evaluate **Judge X** on each of the following five dimensions of judicial performance:

1. **Legal and reasoning ability.** This section asks behavior-based questions about the judge's demonstrated legal decision-making abilities, judgment, and other applied reasoning and analytical skills in court.
2. **Impartiality.** This section includes questions about whether or not the judge bases her or his decisions on complete arguments and evidence from both sides, and whether or not the judge conducts her or himself fairly and objectively, unbiased by outside influences or inherent prejudices.
3. **Professionalism.** This section includes questions about the judge's personal conduct on the bench and level of professionalism as a judge, including but not limited to her or his diligence and attentiveness to the work of the court, degree of preparedness for court, and composure on the bench.
4. **Communication skills.** This section examines the judge's ability to clearly, articulately, and concisely communicate with all participants (e.g., attorneys, litigants, court staff, witnesses, general public) in the courtroom, both orally and in writing.
5. **Management skills.** This section asks questions about the judge's time management skills, court staff management skills, ability to maintain control over the courtroom, and philosophy in handling any courtroom delays or continuances

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

YOUR POSITIVE EXPERIENCES WITH JUDGE X

To help you make a better performance evaluation, please take a few moments to recall some **positive behaviors** you have seen **Judge X** exhibit. Try to remember some positive examples of **Judge X's** behavior along each of the five dimensions of judicial performance. In your direct experience(s) appearing before **Judge X**, what did s/he do well?

Refer to **The Dimensions of Judicial Performance** on page 2 for a description of the five dimensions of judicial performance.

After thinking about these examples of **Judge X's** behavior, please write down any three of the **positive** behaviors you recalled about **Judge X**. Your responses are confidential and will not be used in any way that could identify you as a respondent.

1. _____

2. _____

3. _____

YOUR NEGATIVE EXPERIENCES WITH JUDGE X

Now, please take a few moments to recall some **negative behaviors** you have seen **Judge X** exhibit. Try to remember some negative examples of **Judge X's** behavior along each of the five dimensions of judicial performance. In your direct experience(s) appearing before **Judge X**, what did s/he do poorly?

Refer to **The Dimensions of Judicial Performance** on page 2 for a description of the five dimensions of judicial performance.

After thinking about these examples of **Judge X's** behavior, please write down any three of the **negative** behaviors you recalled about **Judge X**. Your responses are confidential and will not be used in any way that could identify you as a respondent.

1. _____

2. _____

3. _____

JUDICIAL PERFORMANCE EVALUATION INSTRUCTIONS

This Illinois Judicial Performance Evaluation is divided into five sections. You will be asked to focus on and assess **Judge X's** performance on a different dimension of judicial performance (legal & reasoning ability, impartiality, professionalism, communication skills, and management skills) in each of these sections.

CONFIDENTIALITY

Your honest and accurate evaluations of **Judge X** are a crucial component of the judicial performance evaluation program. For this reason, the evaluations you provide are confidential.

All data are stored in a NCSC-sponsored database on a secure server and will not be shared with **Judge X** or any other Illinois official. The summary report reviewed by the judge and project facilitator will contain average ratings, average scores, and frequency distributions of ratings; it will also supply any written evaluation comments you volunteer in each of the subsequent sections. The summary report thus contains no information that would enable the judge to identify you as a respondent, except for any identifying details you choose to include in your written evaluation comments to the judge about his or her performance.

Because your written evaluation comments will be listed verbatim in the summary report for the judge to review, please do not include information in your written evaluation comments that you believe will identify you to the judge.

In the following five sections, please give serious thought to each item. Base your evaluations on your own personal experience in the past 12 months with the judge, and NOT on how you think the judge behaves in other proceedings, others' impressions of the judge, or the judge's reputation. For each item in the evaluation, please consider the context in which you have directly worked with **Judge X**. For those items that you are able to evaluate based on your personal experience with **Judge X**, please indicate the extent to which you witnessed the described behavior using the following scale:

1. I **never** observed the behavior in question.
2. I **rarely** observed the behavior in question.
3. I observed the behavior in question **some** of the time.
4. I **frequently** observed the behavior in question.
5. I observed the behavior in question **every** time.

Please use the N/A response option instead of the above rating scale when: (a) the circumstances under which the described judicial behavior might arise never actually occurred in your experiences with **Judge X**, and/or (b) you do not possess the technical knowledge necessary to evaluate the judge on a particular item.

LEGAL AND REASONING ABILITY

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge's ruling cited the applicable substantive law.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge adhered to the appropriate rules of procedure.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge applied rules of evidence relevant to the case.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge failed to provide a proper legal basis for a decision.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge's decision was inconsistent with rules of evidence.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge drew a conclusion about the case that was inconsistent with the evidence presented.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge incorrectly interpreted a previous court decision applicable to the case.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge correctly interpreted a relevant higher court ruling.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge's decision followed logically from the evidence presented.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge had difficulty applying the law to the facts in the case.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge incorrectly identified the main issues in a case.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge assisted the parties in reaching agreement.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge assisted parties in narrowing key issues in dispute.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge quickly resolved problems that arose during the proceedings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge rendered a well-reasoned decision.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please provide any additional perceptions you have about this judge's *legal and reasoning ability*. You may use this space to identify additional strengths or weaknesses, or to clarify your response(s) to any item in this section. Your comments will be provided to **Judge X** to review for self-improvement purposes.

IMPARTIALITY

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge treated one of the two parties with favoritism.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge weighed the evidence impartially.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge carefully considered arguments from both sides before ruling.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge conducted the proceeding in a neutral manner.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge appeared to decide the outcome of the case before all evidence was presented.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge made a decision before hearing all of the argument.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge unnecessarily restricted an attorney's presentation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge allowed legally irrelevant information to sway her or his decision.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge treated the parties equally.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge's decision was unduly influenced by extralegal factors (e.g., possible public criticism, political interests, fear of appeal, special gifts/favors).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

If you indicated that **Judge X** was influenced by extralegal factors (e.g., possible public criticism, political interests, fear of appeal, special gifts/favors), please select all influencing extralegal factors.

- Possible public criticism
- Political interests
- Fear of appeal
- Special gifts/favors
- Other: _____

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge's decision was unduly influenced by personal characteristics of one of the parties (e.g., race, ethnicity, gender, age, socioeconomic status).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

If you indicated that **Judge X's** decisions were influenced by personal characteristics of one of the parties (e.g., race, ethnicity, gender, age, socioeconomic status), please select all influencing party characteristics.

- Race/Ethnicity
- Gender
- Age
- Socioeconomic status
- Other: _____

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge appeared to maintain an open mind during proceedings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge appeared to categorically favor one side (prosecution or defense) over the other.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please provide any additional perceptions you have about this judge's *impartiality*. You may use this space to identify additional strengths or weaknesses, or to clarify your response(s) to any item in this section. Your comments will be provided to **Judge X** to review for self-improvement purposes.

PROFESSIONALISM

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge appeared to be attending to issues irrelevant to his/her courtroom duties during the court proceeding at hand.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge overreacted to incidents in the courtroom.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge retaliated against an attorney through her or his decision(s) in a case.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge took measures to curb unprofessional attorney behavior during a proceeding.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge disrupted a court proceeding (e.g., with personal cell phone usage, by having discussions with observers, by interjecting personal comments).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge addressed individuals (e.g., attorneys, court staff, litigants, public, witnesses) disrespectfully in the courtroom.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge used unnecessary intimidation to maintain control of the courtroom.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge appeared to be bored with a case.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge carefully reviewed evidentiary materials in the case.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge was unprepared for court.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge maintained a professional demeanor in the courtroom.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please provide any additional perceptions you have about this judge's *professionalism*. You may use this space to identify additional strengths or weaknesses, or to clarify your response(s) to any item in this section. Your comments will be provided to **Judge X** to review for self-improvement purposes.

COMMUNICATION SKILLS

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge issued an order that was clearly written.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge spoke clearly during the court proceeding.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When issuing an oral decision, the judge failed to provide an explanation for the decision.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge gave reasons for a ruling when needed.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge made sure all parties (attorneys and the clients they represent) understood the court proceedings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge's oral communication in court was easily understood.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge contradicted herself or himself during the court proceedings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge issued concise oral decisions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge listened carefully during the court proceeding.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please provide any additional perceptions you have about this judge's *communication skills*. You may use this space to identify additional strengths or weaknesses, or to clarify your response(s) to any item in this section. Your comments will be provided to **Judge X** to review for self-improvement purposes.

MANAGEMENT SKILLS

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge started courtroom proceedings on time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge kept the cases moving promptly.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge maintained an orderly pre-trial schedule.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge lost control of the courtroom.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge failed to explain the reason for a delay.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Considering the amount of case law required to make the decision, the judge was slow in rendering a decision.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
For a continuance, the judge required parties to show cause.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge allowed the appropriate amount of time for each case.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge enforced court deadlines (e.g., for filing documents).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge used courtroom time efficiently.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge provided court staff with clear direction.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please provide any additional perceptions you have about this judge's *management skills*. You may use this space to identify additional strengths or weaknesses, or to clarify your response(s) to any item in this section. Your comments will be provided to **Judge X** to review for self-improvement purposes.

OVERALL EVALUATION

	VERY POOR	POOR	FAIR	GOOD	VERY GOOD
Taking everything into account, how would you rate the performance of this judge?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

YOUR DEMOGRAPHIC INFORMATION (OPTIONAL)

To help us interpret the results of this performance evaluation, please answer a few questions about your background.

CONFIDENTIALITY

Any information you provide on this screen will only be used by NCSC to assess the quality of this evaluation instrument and to minimize bias in the Judicial Performance Evaluation process. Demographic information will not be shared with **Judge X** or any other Illinois official, nor will it be used in any other way that could identify you as a respondent in this Judicial Performance Evaluation.

How many years have you been a practicing attorney? _____ years

How many years have you been a practicing attorney **in the state of Illinois**? _____ years

In which judicial circuit is most of your work conducted? _____ Judicial Circuit

Which of the following options best describe your **primary** area of practice? That is, what area of law comprises the majority of your cases?

- Civil tort - defense
- Civil tort - plaintiff
- Criminal - defense attorney
- Criminal - prosecution
- Commercial/ general civil
- Juvenile offender or dependency
- Domestic relations/Family law
- Estate/Probate
- Traffic
- Government practice
- Other: _____

Which of the following environments best describes your **primary** work setting?

- State's attorney office
- Attorney general's office
- Public defender
- Legal aid
- In-house Corporate Counsel
- Private practice
- Other: _____

Which of the following options best describes your racial/ethnic background? If you identify with two or more of the options below, select "Other" and describe.

(Select one)

- Caucasian/White
- African American/Black
- Hispanic/Latino
- Asian/Pacific Islander
- Native American
- Other: _____

What is your gender? (Select one)

- Male
- Female

Please share any comments you may have about this evaluation, including any suggestions for how it may be improved.

Thank you for your participation in the Illinois Judicial Performance Evaluation of **Judge X**. Please **return this evaluation form** to XXXX XXX by email (XXXX@XXXX.org) or fax (303-XXX-XXXX, ATTN: XXXX XXX) by **END DATE**. Your feedback is appreciated.

If you would like to know more about the Illinois Judicial Performance Evaluation program, please contact the program manager, XX XXXXXXXX XXXX, by phone at (217) XXX-XXXX or by email at XXXX@court.state.il.us.

SUPREME COURT OF ILLINOIS JUDICIAL PERFORMANCE EVALUATION PROGRAM EVALUATION FOR COURT PERSONNEL

You have been nominated by **Judge X** to participate in his/her judicial performance evaluation. This evaluation form contains a series of multiple choice and open-ended questions about **Judge X's** behavior along 5 dimensions of judicial performance that should take no more than 20 minutes to complete.

If you have any questions about this evaluation form, please contact our Help Desk specialist, XXXX XXX, at XXXX@XXXX.org or 303-XXX-XXXX.

For more information about the Illinois Judicial Performance Evaluation Program, please contact the Illinois JPE program manager, XX XXXXX XXXXX, at XXXXX@court.state.il.us or 217-XXX-XXXX.

THE DIMENSIONS OF JUDICIAL PERFORMANCE

In this evaluation form, you will be asked to evaluate **Judge X** on each of the following five dimensions of judicial performance:

1. **Legal and reasoning ability.** This section asks behavior-based questions about the judge's demonstrated legal decision-making abilities, judgment, and other applied reasoning and analytical skills in court.
2. **Impartiality.** This section includes questions about whether or not the judge bases her or his decisions on complete arguments and evidence from both sides, and whether or not the judge conducts her or himself fairly and objectively, unbiased by outside influences or inherent prejudices.
3. **Professionalism.** This section includes questions about the judge's personal conduct on the bench and level of professionalism as a judge, including but not limited to her or his diligence and attentiveness to the work of the court, degree of preparedness for court, and composure on the bench.
4. **Communication skills.** This section examines the judge's ability to clearly, articulately, and concisely communicate with all participants (e.g., attorneys, litigants, court staff, witnesses, general public) in the courtroom, both orally and in writing.
5. **Management skills.** This section asks questions about the judge's time management skills, court staff management skills, ability to maintain control over the courtroom, and philosophy in handling any courtroom delays or continuances.

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

YOUR POSITIVE EXPERIENCES WITH JUDGE X

To help you make a better performance evaluation, please take a few moments to recall some **positive behaviors** you have seen **Judge X** exhibit. Try to remember some positive examples of **Judge X's** behavior along each of the five dimensions of judicial performance. In your direct experience(s) working with **Judge X**, what did s/he do well?

Refer to **The Dimensions of Judicial Performance** on page 1 for a description of the five dimensions of judicial performance.

After thinking about these examples of **Judge X's** behavior, please write down any three of the **positive** behaviors you recalled about **Judge X**. Your responses are confidential and will not be used in any way that could identify you as a respondent.

1. _____

2. _____

3. _____

YOUR NEGATIVE EXPERIENCES WITH JUDGE X

Now, please take a few moments to recall some **negative behaviors** you have seen **Judge X** exhibit. Try to remember some negative examples of **Judge X's** behavior along each of the five dimensions of judicial performance. In your direct experience(s) working with **Judge X**, what did s/he do poorly?

Refer to **The Dimensions of Judicial Performance** on page 1 for a description of the five dimensions of judicial performance.

After thinking about these examples of **Judge X's** behavior, please write down any three of the **negative** behaviors you recalled about **Judge X**. Your responses are confidential and will not be used in any way that could identify you as a respondent.

1. _____

2. _____

3. _____

JUDICIAL PERFORMANCE EVALUATION INSTRUCTIONS

This Illinois Judicial Performance Evaluation is divided into five sections. You will be asked to focus on and assess **Judge X**'s performance on a different dimension of judicial performance (legal & reasoning ability, impartiality, professionalism, communication skills, and management skills) in each of these sections.

CONFIDENTIALITY

Your honest and accurate evaluations of **Judge X** are a crucial component of the judicial performance evaluation program. For this reason, the evaluations you provide are confidential.

All data are stored in a NCSC-sponsored database on a secure server and will not be shared with **Judge X** or any other Illinois official. The summary report reviewed by the judge and project facilitator will contain average ratings, average scores, and frequency distributions of ratings; it will also supply any written evaluation comments you volunteer in each of the subsequent sections. The summary report thus contains no information that would enable the judge to identify you as a respondent, except for any identifying details you choose to include in your written evaluation comments to the judge about his or her performance.

Because your written evaluation comments will be listed verbatim in the summary report for the judge to review, please do not include information in your written evaluation comments that you believe will identify you to the judge.

In the following five sections, please give serious thought to each item. Base your evaluations on your own personal experience in the past 12 months with the judge, and NOT on how you think the judge behaves in other proceedings, others' impressions of the judge, or the judge's reputation. For each item in the evaluation, please consider the context in which you have directly worked with **Judge X**. For those items that you are able to evaluate based on your personal experience with **Judge X**, please indicate the extent to which you witnessed the described behavior using the following scale:

1. I **never** observed the behavior in question.
2. I **rarely** observed the behavior in question.
3. I observed the behavior in question **some** of the time.
4. I **frequently** observed the behavior in question.
5. I observed the behavior in question **every** time.

Please use the N/A response option instead of the above rating scale when: (a) the circumstances under which the described judicial behavior might arise never actually occurred in your experiences with **Judge X**, and/or (b) you do not possess the technical knowledge necessary to evaluate the judge on a particular item.

LEGAL AND REASONING ABILITY

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge incorrectly identified the main issues in a case.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge assisted the parties in reaching agreement.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge quickly resolved problems that arose during the proceedings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge assisted parties in narrowing key issues in dispute.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please provide any additional perceptions you have about this judge's *legal and reasoning ability*. You may use this space to identify additional strengths or weaknesses, or to clarify your response(s) to any item in this section. Your comments will be provided to **Judge X** to review for self-improvement purposes.

IMPARTIALITY

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge treated one of the two parties with favoritism.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge conducted the proceeding in a neutral manner.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge appeared to decide the outcome of the case before all evidence was presented.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge unnecessarily restricted an attorney's presentation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge treated the parties equally.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge's decision was unduly influenced by extralegal factors (e.g., possible public criticism, political interests, fear of appeal, special gifts/favors).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

If you indicated that **Judge X** was influenced by extralegal factors (e.g., possible public criticism, political interests, fear of appeal, special gifts/favors), please select all influencing extralegal factors.

- Possible public criticism
- Political interests
- Fear of appeal
- Special gifts/favors
- Other: _____

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge's decision was unduly influenced by personal characteristics of one of the parties (e.g., race, ethnicity, gender, age, socioeconomic status).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

If you indicated that **Judge X's** decisions were influenced by personal characteristics of one of the parties (e.g., race, ethnicity, gender, age, socioeconomic status), please select all influencing party characteristics.

- Race/Ethnicity
- Gender
- Age
- Socioeconomic status
- Other: _____

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge appeared to maintain an open mind during proceedings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge appeared to categorically favor one side (prosecution or defense) over the other.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please provide any additional perceptions you have about this judge's *impartiality*. You may use this space to identify additional strengths or weaknesses, or to clarify your response(s) to any item in this section. Your comments will be provided to **Judge X** to review for self-improvement purposes.

PROFESSIONALISM

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge appeared to be attending to issues irrelevant to his/her courtroom duties during the court proceeding at hand.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge overreacted to incidents in the courtroom.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge took measures to curb unprofessional attorney behavior during a proceeding.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge disrupted a court proceeding (e.g., with personal cell phone usage, by having discussions with observers, by interjecting personal comments).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge addressed individuals (e.g., attorneys, court staff, litigants, public, witnesses) respectfully in the courtroom.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge used unnecessary intimidation to maintain control of the courtroom.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge treated court employees respectfully regardless of position.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge appeared to be bored with a case.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge was unprepared for court.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge maintained a professional demeanor in the courtroom.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please provide any additional perceptions you have about this judge's *professionalism*. You may use this space to identify additional strengths or weaknesses, or to clarify your response(s) to any item in this section. Your comments will be provided to **Judge X** to review for self-improvement purposes.

COMMUNICATION SKILLS

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge spoke clearly during the court proceeding.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When issuing an oral decision, the judge failed to provide an explanation for the decision.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge gave reasons for a ruling when needed.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge made sure all parties (attorneys and the clients they represent) understood the court proceedings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge's oral communication in court was easily understood.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge listened carefully during the court proceedings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please provide any additional perceptions you have about this judge's *communication skills*. You may use this space to identify additional strengths or weaknesses, or to clarify your response(s) to any item in this section. Your comments will be provided to **Judge X** to review for self-improvement purposes.

MANAGEMENT SKILLS

	(1) NEVER	(2) RARELY	(3) SOMETIMES	(4) FREQUENTLY	(5) EVERY TIME	N/A
The judge started courtroom proceedings on time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge kept the cases moving promptly.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge maintained an orderly pre-trial schedule.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge lost control over the courtroom.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge failed to explain the reason for a delay.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
For a continuance, the judge required parties to show cause.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge allowed the appropriate amount of time for each case.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge enforced court deadlines (e.g., for filing documents).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge used courtroom time efficiently.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge provided court staff with clear direction.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The judge thoroughly addressed concerns raised by court staff.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please provide any additional perceptions you have about this judge's *management skills*. You may use this space to identify additional strengths or weaknesses, or to clarify your response(s) to any item in this section. Your comments will be provided to **Judge X** to review for self-improvement purposes.

OVERALL EVALUATION

	VERY POOR	POOR	FAIR	GOOD	VERY GOOD
Taking everything into account, how would you rate the performance of this judge?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

YOUR DEMOGRAPHIC INFORMATION (OPTIONAL)

To help us interpret the results of this performance evaluation, please answer a few questions about your background.

CONFIDENTIALITY
Any information you provide on this screen will only be used by the NCSC to assess the quality of this evaluation instrument and to minimize bias in the Judicial Performance Evaluation process. Demographic information will not be shared with **Judge X** or any other Illinois official, nor will it be used in any way that could identify you as a respondent in this Judicial Performance Evaluation.

Which of the following options best describes your racial/ethnic background? If you identify with two or more of the options below, select "Other" and describe.

(Select one)

- Caucasian/White
- African American/Black
- Hispanic/Latino
- Asian/Pacific Islander
- Native American
- Other: _____

What is your gender? (Select one)

- Male
- Female

Please share any comments you may have about this evaluation, including any suggestions for how it may be improved.

Thank you for your participation in the Illinois Judicial Performance Evaluation of **Judge X**. Please **return this evaluation form** to XXXX XXX by email (XXXX@XXXX.org) or fax (303-XXX-XXXX, ATTN: XXXX XXX) **by END DATE**. Your feedback is appreciated.

If you would like to know more about the Illinois Judicial Performance Evaluation program, please contact the program manager, XX XXXXXXXX XXXX, by phone at (217) XXX-XXXX or by email at XXXX@court.state.il.us.



INSTITUTE *for the* ADVANCEMENT
of the AMERICAN LEGAL SYSTEM



UNIVERSITY *of*
DENVER

Institute for the Advancement of the American Legal System

University of Denver

John Moyer Hall, 2060 South Gaylord Way

Denver, CO 80208

Phone: 303.871.6600 <http://iaals.du.edu>