

The Courtroom-Observation Program of the Utah Judicial Performance Evaluation Commission

Nicholas H. Woolf & Jennifer MJ Yim

The State of Utah has recently introduced several innovations to its judicial performance evaluation program. Since the 1960s, such programs have been used in a growing number of states to inform decisions about judicial retention and to provide feedback to judges about their performance. The evaluations have traditionally been limited to surveys completed by a variety of stakeholders in the court system. But several states have begun to expand the scope of their evaluation programs beyond surveys, and one of these innovations is to observe judges in their courtrooms.

Utah has developed its courtroom-observation program to become a major component of its overall judicial performance evaluation. This article describes the history of judicial performance evaluations in Utah, the introduction of courtroom observation by lay observers, and the recent innovations to the observation program, including its focus on procedural fairness, use of qualitative evaluation methods, and use of systematic content analysis of the observers' narrative reports.

BACKGROUND TO THE UTAH COURTROOM-OBSERVATION PROGRAM

History of Judicial Performance Evaluation

Judicial performance evaluation (JPE) has steadily grown in popularity since the first program in Alaska in 1967.¹ In 1985, the American Bar Association issued its first guidelines for JPE, and by 1998 "approximately 23 states had some sort of JPE program in place or under development."² In 2005, the American Bar Association updated its JPE guidelines, proposing that all court systems implement a formal program to promote judicial self-improvement, enhance the quality of the judiciary, and pro-

vide relevant information to voters. These guidelines provide evaluative criteria in five areas (legal ability, integrity and impartiality, communication skills, professionalism and temperament, and administrative capacity), and suggest various aspects of evaluation design, including evaluating actual behaviors rather than the general qualities of judges, collecting data from multiple sources, and using experts for developing evaluation methods and for collecting and analyzing data.³

It is not clear that these guidelines have uniformly been followed, and David Brody has proposed a nationwide assessment of the processes currently in use in JPE programs.⁴ One issue seems clear from the JPE literature: reliance on surveys that result in a single metric of performance does not promote acceptance and trust of JPE by judges.⁵ While some studies of JPE programs find that judges have predominantly positive comments about JPE, including the usefulness of feedback and the appropriateness of criteria,⁶ other studies document judicial mistrust of survey-based JPE. In a Colorado study, while judges reported positive comments about JPE, many criticized the survey methodology used as "unscientific" and "inherently flawed."⁷ Another study found that some JPE surveys appear to be "a fishing expedition trolling about for evidence of judicial misbehavior or performance problems"⁸ rather than an evaluation of clear performance standards. Analyses of survey data show that respondents tend to evaluate female and minority judges more harshly⁹ and that aggregating scores across a wide variety of criteria, or using a single measure for different kinds of courts, reduces the validity of the survey results.¹⁰ Concerns with survey methodologies provide one explanation for the slow adoption of JPE.¹¹

The authors would like to thank David Roth, Joanne Slotnik, and David Turner for their constructive feedback on earlier drafts of this article.

Footnotes

1. Richard C. Kearney, *Judicial Performance Evaluation in the States*, 22 PUB. ADMIN. Q. 468 (Winter 1999).
2. *Id.* at 471.
3. AMERICAN BAR ASSOCIATION, JUDICIAL DIVISION, LAWYERS CONFERENCE, ABA BLACK LETTER GUIDELINES FOR JUDICIAL PERFORMANCE EVALUATION (Feb. 2005), available at http://www.americanbar.org/groups/judicial/conferences/lawyers_conference/resources/judicial_performance_resources.html (last visited Feb. 24, 2012).
4. David C. Brody, *The Use of Judicial Performance Evaluation to Enhance Judicial Accountability, Judicial Independence, and Public Trust*, 86 DENVER U. L. REV. 115 (2008).
5. *Id.* at 155-56.
6. E.g., AMERICAN JUDICATURE SOCIETY—WASHINGTON CHAPTER, COMMITTEE ON JUDICIAL PERFORMANCE STANDARDS AND EVALUATION,

- FINAL REPORT OF THE WASHINGTON STATE JUDICIAL PERFORMANCE EVALUATION PILOT PROJECT 29 (2002); KEVIN M. ESTERLING & KATHELEEN M. SAMPSON, JUDICIAL RETENTION EVALUATION PROGRAMS IN FOUR STATES: A REPORT WITH RECOMMENDATIONS 39 (American Judicature Society 1998).
7. Brody, *supra* note 4, at 147.
8. Kearney, *supra* note 1, at 484.
9. Rebecca Wood, Sylvia R. Lazos & Mallory Waters, Sacrificing Diversity for "Quality": How Judicial Performance Evaluations Are Failing Women & Minorities, *Scholarly Works*, Paper 24 (Apr. 5, 2010), <http://scholars.law.unlv.edu/facpub/24/>.
10. Rebecca Gill & Sylvia R. Lazos, Reflections in Response to the Nevada Judicial Evaluation Pilot Project, Working Paper Series, Abstract No. 1539825 (Dec. 18, 2009), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1539825.
11. Seth S. Andersen, *Judicial Retention Evaluation Programs*, 34 LOY. L.A. L. REV. 1375 (2001).

In response to these concerns, some states have broadened their JPE programs to include reviews of case-management records, solicitation of public comments, or interviews with judges.¹² Several states also have courtroom-observation programs of various kinds. In Alaska, an independent organization trains volunteers to observe courtroom proceedings, complete surveys, and write narrative comments, a selection of which are reproduced in a report to the Alaska Judicial Council and included as input to the JPE process.¹³ Citizen organizations in both New York and Minnesota, neither of which have state JPE programs, train volunteers to conduct courtroom monitoring as outsiders to the legal system; these organizations publicize their reports and recommendations to the public and the court system.¹⁴ Colorado has multiple JPE commissions by geographic jurisdiction, and courtroom observation is conducted by the JPE commission members themselves, who receive training in collecting various sources of data for their evaluations. In Nevada, a 2010 ballot measure to establish a comprehensive JPE program included a formal courtroom-observation program, but the measure was defeated.

History of the Utah Courtroom-Observation Program

Since 1986, the Utah State Courts have conducted JPEs for state-court judges. However, in 2008, the Utah State Legislature passed the Judicial Performance Evaluation Commission Act,¹⁵ altering the judicial-retention process by moving responsibility for JPE from Utah's judicial branch to a newly created, independent Judicial Performance Evaluation Commission (JPEC). The statute was based on the work of the Judicial Retention Election Task Force, which was composed of legislators and judges. The task force met for approximately six months to produce draft legislation, receiving input from the Institute for the Advancement of the American Legal System at the University of Denver regarding JPE efforts across the nation.

By statute, Utah's JPEC is composed of 13 volunteer members, four appointed by each branch of government, and a final member who is the executive director of the Governor's Commission on Criminal and Juvenile Justice. The statute sets members' terms, limits the number of attorneys, balances political-party affiliations, precludes sitting legislators and sitting judges from membership, and creates an executive-director position to staff the commission. The statute covers all state, county, and municipal judges (county and municipal

courts are known collectively in Utah as justice courts). Utah judges are appointed either by the state's governor, county government, or a municipal appointing authority, and all face uncontested retention elections at the conclusion of their terms. Two important purposes of the new JPEC were to provide Utah voters with meaningful and independent information about judges and to provide for retention recommendations from JPEC during retention elections.

The statute requires JPEC to complete a JPE twice during each six-year term of office for all sitting judges: once at midterm and again near the end of term in preparation for the retention election.¹⁶ Each JPE must include a judicial performance survey, courtroom observation, and a review of judicial disciplinary records. The statute sets minimum performance standards, makes requirements about the publication and availability of evaluation data, and specifies required survey respondent groups and survey categories.

As noted, the statute mandates the creation of a courtroom-observation program, and requires JPEC to "make rules concerning the conduct of courtroom observation" with respect to who may perform observations, whether they are to be completed in person or electronically, and the principles and standards used to evaluate the observed behaviors.¹⁷ While the statute is specific about the details of the survey design and implementation, the legislature required JPEC to determine both the details of the courtroom-observation program and how the resulting information should be used in formulating judicial-retention recommendations.

In its first effort, JPEC developed an observation survey instrument similar to the survey instrument completed by attorneys. JPEC recruited and trained volunteer observers who observed judges in court, scored judges on the survey questions, and added comments as they thought appropriate. Results closely mimicked the survey results and generated no new information. Additionally, JPEC members worried that an observation survey based on a small sample size of about five observers per judge was fraught with reliability concerns.

To glean different information from that available in its surveys, JPEC changed its approach in two major ways. First, the courtroom-observation program focused exclusively on procedural fairness.¹⁸ Second, the observation instrument elicited qualitative rather than quantitative information.

The focus on procedural fairness recognized the well-established yet counterintuitive body of social-science research that

12. Rebecca Love Kourlis & Jordan M. Singer, *Using Judicial Performance Evaluations to Promote Judicial Accountability*, 90 JUDICATURE 200 (2007).

13. For the most recent report of the independent citizen organization, Alaska Judicial Observers, Inc., on its in-court observations of Alaska judges, see Alaska Judicial Observers, 2010 Biennial Report, <http://www.ajc.state.ak.us/retention/retent2010/judobse10.pdf>.

14. The Fund for Modern Courts in New York has been training volunteers to observe judges and publicly report their observations since 1975. Recent reports may be found at <http://www.moderncourts.org/Programs/monitoring.html>. A group in Minnesota, called WATCH, indicates on its website that it has trained volunteers to observe court hearings in Hennepin County, Minnesota.

See <http://www.watchmn.org/>.

15. 2008 Utah Laws ch. 248, § 5, *codified at* Utah Code Ann. 78A-12-201.

16. Utah Supreme Court judges serve ten-year terms and so receive two midterm evaluations and a retention-election evaluation. The courtroom-observation component is currently administered for all Utah judges except appellate level judges and part-time justice-court judges.

17. 2008 Utah Laws ch. 248, § 7, *codified at* Utah Code Ann. 78A-12-203(3).

18. We consider the terms *procedural justice* and *procedural fairness* to be synonymous, and for consistency use the term procedural fairness throughout.

has developed over the last 20 years.¹⁹ The central conclusion of this research is that people are strongly influenced by their judgments of how fairly they are treated by authorities of all kinds, including courts and judges. Their trust in the law and their voluntary acceptance of judges' decisions are more strongly influenced by their judgment of the fairness of the procedures and their treatment than by the favorability of the outcome they receive in court. The research does not in any way suggest that the desirability of achieving a party's outcome is not of great concern, but rather emphasizes the particular importance people place upon fair procedures and fair treatment. The behavior of judges has a central role in producing perceptions of procedural fairness. Consequently, JPEC decided to focus the courtroom-observation portion of the JPE on this important area.

The decision to adopt qualitative methods was important too. Quantitative methods, such as those used in surveys, produce information in different ways from qualitative methods, and the two methods complement one another to generate a more complete picture of what is being evaluated. Surveys collect structured information from respondents about clearly defined descriptions of behaviors (e.g., *the judge properly applies the rules of civil procedure*), and measure responses in statistically valid ways. In contrast, qualitative evaluations seek unstructured information about behavior in its natural context (e.g., the narrative comments about judicial behaviors written by courtroom observers), without specifying in advance exactly what those behaviors might be. JPEC decided that because context was so important to evaluating procedural-fairness behaviors, qualitative methods would be most appropriate for the courtroom-observation program.

To provide a systematic analysis of the qualitative courtroom-observation data, and to produce careful summaries of the large amount of information that would be produced by observers, JPEC hired an outside consultant with expertise in qualitative-evaluation research to conduct a pilot study of five judges.²⁰ For each judge in the pilot study, the consultant prepared a content analysis of the judge's courtroom-observer narratives and a summary report presenting the analysis in a concise format.²¹ After completion, JPEC staff and the Courtroom Observation Subcommittee chairperson visited each of the five judges to solicit feedback both on the results of the pilot study and on the observation program more generally. JPEC concluded from the positive responses to the pilot study that the summary reports were effective in enhancing the credibility of the courtroom-observation data and so decided to incorporate content analyses into the program.²²

JPEC consequently adopted administrative rules to codify the details of the courtroom-observation program,²³ including the selection and terms of service of observers, the required training, the numbers of observations per judge, and the minimum time spent observing each courtroom. The rule also established procedural-fairness principles as the basis for the observers' evaluations.

THE CURRENT COURTROOM-OBSERVATION PROGRAM

The courtroom-observation program is coordinated by a part-time JPEC staff person, and overseen by JPEC's advisory Courtroom Observation Subcommittee.²⁴ This section describes the recruitment and training of observers, the observation procedures, and the content analysis and summarization of the observation data.

Observer Recruitment and Training

Recruitment. JPEC seeks volunteers with a "broad and varied range of life experiences" who can commit to a one-year renewable term of service.²⁵ The coordinator recruits volunteers using a statewide advertising and public-outreach effort that taps local media, continuing-education programs, and government and nonprofit organizations. JPEC staff and commission members provide an orientation, and screen applicants based on written applications and an interview process. JPEC excludes those lacking basic computer skills or access, those involved in pending litigation in the state, and convicted felons. Those with professional involvement in the courts and certain types of relationships with judges or the courts are also automatically ineligible. During 2011, JPEC had 28 volunteer courtroom observers who produced 330 courtroom-observation reports for 65 judges.

Training. Over time, JPEC has developed a thorough training program for volunteer observers. Initial training includes an orientation about the overall evaluation process of which courtroom observation is a part; a primer on levels of court and court process; confidentiality, nondisclosure, and conflict-of-interest requirements; an introduction to the principles and standards of procedural fairness; and technical instruction on the data-collection process, including proper use of the courtroom-observation instrument. Some sections are taught by commission members, some by JPEC staff, and some by senior judges or court administrators. Initial training includes practice observations and mentoring opportunities with experienced observers. Periodic in-service trainings and other meetings, such as Ask-a-Judge sessions with senior judges and events with commission members, help to ensure ongoing effective observations and volunteer retention.

19. For a comprehensive discussion of the importance of procedural fairness, see the special issue of *Court Review*, comprising the first two issues of Volume 44 (2007-08), which is available on the web at <http://aja.ncsc.dni.us/courtrv/cr44-1/CR44-1-2.pdf>.

20. The first author of this article, Nicholas H. Woolf, is the qualitative-research consultant who conducted the pilot study and is conducting the 2012-2014 content analyses and preparation of summary reports.

21. Further information about the content analysis and summary

report appears below in the section, "Current Program."

22. Further information about feedback from the five pilot study judges to their summary reports appears below in the section, "Uses of Courtroom-Observation Reports."

23. Utah Admin. Code R597-3-3.

24. The second author of this article, Jennifer MJ Yim, is the vice-chairperson of JPEC, appointed by the governor, and is the past chairperson of the Courtroom Observation Subcommittee.

25. Utah Admin. Code R597-3-3.

Observation Procedures

By administrative rule, a minimum of four different volunteers observe each judge for both the midterm and end-of-term evaluations. To date, JPEC has provided almost all judges with five observations, both at midterm and end of term. Each observer observes a judge in person for a minimum of two hours while the judge is actively on the bench. Observations may be completed in one sitting or over several courtroom visits, after which the observer completes the data-collection instrument and submits it electronically to JPEC staff. The identity of the observer is withheld from the judge, JPEC commissioners, and the public. However, JPEC staff records information about who conducted the observation (by observer code), the dates and type of proceedings observed, and the gender of the observer.

The data-collection instrument is narrative-based, instructing the observer to report his or her experiences in the courtroom with respect to the procedural-fairness principles of neutrality, respect, and voice, and to report other notable aspects of the observation experience, including whether they found the judge to be trustworthy, whether they would feel comfortable appearing before the judge as a litigant, and whether the judge was aware of the observation. Trainers instruct observers to write detailed, contextually specific narratives that articulate: (1) the behaviors they observed, and (2) the observer's personal reaction to those behaviors. Observers report that compiling their notes into comprehensive narratives requires several hours of work beyond the time spent in the courtroom.

Content Analysis and Summarization of Observation Data

Criteria. The consultant first developed a set of evaluative criteria for analyzing the observers' narratives. These criteria describe the judicial behaviors that lead people to believe they have been treated fairly in the courtroom. The criteria are based on the "four key procedural fairness principles,"²⁶ *i.e.*, whether recipients feel that they have been given *voice*, have been treated with *neutrality* and *respect*, and experience *trust* in the judge and the legal system. The consultant reviewed the procedural-fairness literature and identified approximately 50 judicial behaviors that have been proposed to lead those in the courtroom to experience (or not experience) procedural fairness. These were clustered into 20 criteria, which were

grouped into six approximately equal-sized groups. Table 1 displays the six groups of 20 criteria, with brief descriptions of each criterion.²⁷

Content analysis. A content analysis describes the content of narrative data, but goes beyond simple summarizing of information. The consultant developed a set of 29 "codes," or categories, consisting of the 20 evaluative criteria (see Table 1) and nine additional codes that provide greater detail for some of the criteria (for example, the criterion *courtroom tone & atmosphere* was divided into two for purposes of coding: *general demeanor* and *courtroom tone*). Each code has a definition to facilitate consistent coding of observer comments.

The observers' narratives are segmented into "codable units" that reflect a single experience or reaction to a judicial behavior. These range from a single phrase for a straightforward observation to one or more paragraphs describing in detail something that occurred in court. Each codable unit is then "coded," *i.e.*, associated with one or more codes, using the qualitative-analysis software program Atlas.ti. Observers are encouraged to write in whatever manner best reflects their experience of the judge's behaviors, and codes are selected to best describe the observers' reactions to judicial behaviors, rather than to interpret the fairness of the judge's behavior in any objective sense. In this way the evaluation assesses how judicial behaviors are experienced by others.

Summary report. After coding all five observers' narratives for a judge, the coded units for each criterion are analyzed for themes and commonalities, and then synthesized into one or more short paragraphs that evoke as accurately as possible the experience of reading all the narrative associated with each criterion. The paragraphs for all codes are then compiled into a two- to three-page report. The report begins with an "executive summary" that further synthesizes the information into three standard categories: *widely agreed-upon themes*, *minority observations*,²⁸ and *anomalous comments*. Table 2 displays the content and purpose of each of the three summary sections in more detail.²⁹

Uses of Courtroom-Observation Reports

The courtroom-observation program has three primary audiences: the general public, judges, and JPEC. Each stakeholder group uses and benefits from the reports in different ways.

26. Tom R. Tyler, *Procedural Justice and the Courts*, 44 CT. REV. 30 (2007-08).

27. While trust is one of the four principles of procedural fairness, there is no group of criteria called trust, nor any individual criterion called trust. From the perspective of evaluating performance behaviors, we consider trust an outcome of respect, neutrality and voice, rather than a behavior in itself. Some observers reported difficulty in assessing behaviors themselves as "trustworthy." Similarly, researchers in the management field listed behaviors that engendered employee trust, rather than describing trustworthiness as a behavior itself, see Ellen M. Whitener, et al., *Managers as Initiators of Trust: An Exchange Relationship Framework for Understanding Managerial Trustworthy Behavior*, 23 ACAD. MGMT. REV. (1998). We foresee difficulties with feedback to judges if trust is considered a criterion. It may be one thing to say to a judge, "Your behaviors, such as lack of eye contact, were perceived as not

respectful, and respectful behaviors engender trust," but quite another to say, "Your behaviors were not trustworthy."

28. Minority in this context refers to comments from two, or possibly three, of the five observers, and does not refer to observers' ethnic or racial membership.

29. The courtroom-observation instrument initially collected a numerical rating of the judge on a five-point scale, similar to JPEC's survey. Observers reported difficulty accurately assigning scores in the context of writing contextual narratives. Commissioners felt concern that the scoring did not meaningfully reflect the content of the narrative data or qualitative synthesis, and the qualitative-research consultant felt concern that the act of completing the numerical scores would have a priming effect on observers when writing their narrative comments. Judges in the pilot study were not averse to the removal of the scores. Consequently, JPEC removed the scoring questions from the observation instrument.

TABLE 1: TWENTY EVALUATIVE CRITERIA

CRITERIA	DESCRIPTION
GROUP 1: RESPECTFUL BEHAVIORS	
<i>LISTENING AND FOCUS</i>	Here listening refers to indications of attention and engagement through active listening.
<i>WELL PREPARED AND EFFICIENT</i>	Here efficiency refers to the judge's behaviors. The court's efficiency appears below under "Courtroom tone & atmosphere."
<i>RESPECT FOR OTHER'S TIME</i>	This includes the starting time of sessions as well as all interactions with those in court that take into consideration the value of their time.
<i>RESPECTFUL BEHAVIORS</i>	This refers to specifically described behaviors directed at specific individuals or situations that indicate respect for a person's value or status.
GROUP 2: RESPECTFUL TONE	
<i>COURTESY, POLITENESS, AND PATIENCE</i>	This refers to descriptions of respectful behaviors more generally, not in the context of a specific individual or situation.
<i>COURTROOM TONE AND ATMOSPHERE</i>	This refers more generally to the tone of the court, and includes both the judge's general demeanor over and above specifically respectful behaviors, as well as the atmosphere of the courtroom.
<i>BODY LANGUAGE</i>	This refers to eye contact and facial expressions, general body language, and engaged behavior.
<i>VOICE QUALITY</i>	This refers to both mechanical qualities such as pitch and volume, and emotional qualities such as inexpressive, sarcastic, or exasperated tone.
GROUP 3: NEUTRALITY	
<i>CONSISTENT AND EQUAL TREATMENT</i>	This refers to listening to all sides and treating individuals in similar situations similarly.
<i>ACTS WITH CONCERN FOR INDIVIDUAL NEEDS</i>	This refers to concern for individual differences being taken account of in the judge's actions.
<i>EXPRESSES CONCERN FOR THE INDIVIDUAL</i>	This refers to expressions of concern and due regard for the individual's specific situation, over and above expressions of respect that are included in "Respectful behaviors."
<i>UNHURRIED AND CAREFUL</i>	This refers to allowing sufficient time for the judge and those in court to conduct themselves in a thorough manner.
GROUP 4: VOICE	
<i>CONSIDERED VOICE</i>	This refers both to allowing those in court to express themselves and to the judge's consideration of what was expressed in his/her statements or decision.
<i>FORMAL VOICE</i>	This refers to giving voice based on required procedure without apparent consideration by the judge of what was expressed.
GROUP 5: COMMUNICATION	
<i>COMMUNICATES CLEARLY</i>	This refers both to clarity of speech and to the use of language appropriate to the listener.
<i>ENSURES INFORMATION UNDERSTOOD</i>	This refers to active attention by the judge in ensuring those in court understand all information relevant to them, and includes translation and comprehension for non-native English speakers.
<i>PROVIDES ADEQUATE EXPLANATIONS</i>	This refers to providing sufficient explanation of the basis of decisions and of legal procedure and terminology to ensure that those in court understand proceedings relevant to them.
GROUP 6: APPARENT FAIRNESS	
<i>VENEER OF FAIRNESS</i>	This refers to behaviors intended to demonstrate fairness but backfire due to perceived insincerity.
<i>INEFFECTIVE INGRATIATION</i>	This refers to efforts to be liked, either defensively through apologies, self-deprecation, etc., or assertively through praise or flattery, that backfire due to inappropriateness or perceived insincerity.
<i>MANIPULATION</i>	This refers to behaviors intended to produce feelings of fairness that backfire because the behaviors are perceived to be manipulative.

TABLE 2: OVERVIEW SECTIONS OF THE CONTENT-ANALYSIS SUMMARY REPORT

WIDELY AGREED-UPON THEMES	<p>Behaviors reported by all (or almost all) observers and are thus well established. Deficits mentioned here were widely reported and therefore merit attention.</p> <p>The first statement in this section is an overall summary of the entire set of observer reports.</p> <p>The subsequent statements are not intended to be a complete summary of the observers' comments, but rather highlight the most frequently noted and forcefully expressed themes in the way that the observers expressed them, with the goal of evoking an overall sense of the entire set of observer comments.</p> <p>The final statement in this section indicates the number of observers indicating that they would feel comfortable appearing before the judge.</p>
MINORITY OBSERVATIONS	<p>Behaviors noted by two (or possibly three) observers that would be worth building on (if desirable) or otherwise thinking about avoiding.</p> <p>Not every behavior reported by a minority of observers is summarized here, only those that reflect a notable or somewhat discrepant perspective that was not widely agreed upon.</p>
ANOMALOUS COMMENTS	<p>Comments of one (or in rare cases two) observers that reflect a markedly different or contradictory perspective from all other observers. They are intended to stimulate reflection, such as: why were these observers affected by this behavior; or, does this particular situation tend to lead to this uncharacteristic behavior?</p> <p>Not every anomalous comment is included in this overview section. While all are included in the main body of the report, some are not included in this overview section because they are too minor or appear to reflect something anomalous about the observer rather than the judge.</p>

General public. Recent research suggests that voters in judicial elections are more interested in judges “as guarantors of procedural fairness” than in assessing the individual policy preferences, case outcomes, and partisanship of judges.³⁰ JPEC’s courtroom-observation program addresses this public concern by providing information to voters about each judge’s performance with respect to procedural fairness.

The program also benefits the public in two other ways. First, it provides certain members of the public—the courtroom observers themselves—an important voice in the evaluation process, as proxies for the many voters who never get to court. Second, the program will help to increase procedural fairness by providing feedback to judges about their behaviors. The courtroom-observation summary report provides context-specific and rich information about how judicial behaviors are perceived by observers in terms of procedural fairness. JPEC anticipates that this detailed and candid feedback to judges will help judges make any needed behavioral changes.

In summary, JPEC anticipates that through these three benefits to the public—providing voters with valued procedural-fairness information, providing a way for members of the public, as observers, to have a voice in JPE, and through improvements to procedural fairness through context-specific feedback to judges—the courtroom-observation program will contribute to increased levels of public trust and confidence in the court system in Utah.

Judges. Judges benefit from the courtroom-observation program both at an individual level, through self-improvement, as well as at the level of the overall judiciary. One important goal of the JPEC statute, as well as of the American Bar Association in

its JPE guidelines, is judicial self-improvement. The two evaluations completed during each six-year retention cycle³¹ are similar in content but have different purposes. The midterm JPE is intended to provide feedback for self-improvement, while the end-of-term JPE is intended to provide JPEC with information for its retention-election recommendation. The way judges use their performance evaluations, however, is more continuous. As most judges in Utah are retained, every JPE, whether completed at midterm or end of term, contains valuable self-improvement feedback.³²

For any judge, obtaining specific feedback about courtroom conduct is generally a challenge. Typically, everyone in a courtroom has something at stake or is otherwise involved in the proceedings. Most people would demur if asked by the judge for critical feedback. Courtroom observers are perhaps the only people in court with the unique potential to provide judges with candid and constructive feedback about their procedural-fairness behaviors.

Adopting procedural fairness as the basis for the courtroom-observation program has raised awareness about procedural fairness in Utah’s judiciary. Using procedural fairness as a component of JPE has prompted judges to become more conscious of their courtroom conduct. Indeed, the judiciary has already begun including procedural fairness as a topic in its judicial-education curriculum.

In summary, the procedural-fairness-based courtroom-observation program serves judges in two ways. First, it assists with judicial self-improvement by providing judges with candid feedback based on actual behaviors observed. Second, it has drawn attention to the importance of procedural fairness both for individual judges and for the judiciary as a whole,

30. Jordan M. Singer, *The Mind of The Judicial Voter*, MICH. ST. L. REV., forthcoming, <http://ssrn.com/abstract=1937742> (Oct. 3, 2011).

31. This is the cycle for judges other than justices of the Utah Supreme Court and part-time justice-court judges. See n. 15 *supra*.

32. JPEC provides both the midterm and end-of-term courtroom-observation reports to judges using the same procedures and reporting format.

leading to an increased number of judicial training opportunities on procedural fairness.

JPEC. JPEC benefits from the courtroom-observation program because it provides an additional and different source of data, thus enhancing the credibility of the evaluation for each judge. In addition to the benefits JPEC receives, it also faces a challenge in terms of how it will utilize the courtroom-observation data in the JPE.

With the addition of courtroom-observation data, JPEC's retention recommendation will be based both on contextually rich and systematically analyzed qualitative data from the courtroom observations, and on statistically valid quantitative data from the surveys. Using these multiple types of data increases the quality of the JPE for Utah judges.³³

Through the courtroom-observation program, JPEC has developed a way to systematically assess judicial behaviors in terms of procedural fairness.³⁴ Although JPEC was not statutorily required to evaluate judges on procedural-fairness criteria, the growing body of research demonstrates its particular importance to voters and court users. Indeed, after much deliberation, JPEC decided that the relevance of procedural fairness reached beyond courtroom observation to the full JPE. JPEC therefore unanimously adopted an administrative rule creating a minimum performance standard for procedural fairness. The new minimum performance standard states that the judge's performance must "demonstrate by a preponderance of the evidence, based on courtroom observations and relevant survey responses, that the judge's conduct in court promotes procedural fairness for court participants."³⁵

While JPEC has now promulgated a minimum performance standard governing procedural fairness, it must still decide how courtroom-observation results will be used in the JPE process. JPEC will need to integrate the qualitative courtroom-observation data and the quantitative survey data to decide whether a judge has met the minimum performance standard for procedural fairness. While the three statutory survey categories—judicial temperament & integrity, legal ability, and administrative performance—do not explicitly include procedural fairness, the survey includes questions related to several aspects of procedural fairness. Deciding whether a judge has met the minimum performance standard for procedural fairness will therefore be based on both courtroom observation and relevant survey information. JPEC will determine the relative weight to be given to survey and courtroom-observation findings.

In summary, JPEC benefits from the courtroom-observation program by increasing the quality of the JPE through the addition of qualitative data on procedural fairness. JPEC now faces the challenge of integrating the courtroom-observation results and the survey results to determine whether judges

meet the minimum performance standard established for procedural justice.

FUTURE PLANS

Development of Courtroom-Observation Program

JPEC plans several developments for its courtroom-observation program, in part based on the experience of completing courtroom observations for 72 judges. These plans include updating the observer training program, efforts to increase the diversity of observers, and extending the courtroom-observation program to part-time justice-court judges. Two other potential uses of the program data are also described below.

Training program. Observers write their comments in very different styles. Because the goal is to learn how observers experience the judge's procedural fairness behaviors, no constraints are put on the way in which the observers most comfortably express themselves. However, because the role of the observer is intended as a proxy for courtroom participants, JPEC trains observers to simply describe how they experience the judge's procedural fairness behaviors, not to interpret their feelings or the judge's behaviors in terms of their personal knowledge of procedural fairness or other social-science principles. To assist ongoing training, the consultant periodically provides examples to JPEC of ideal comments (containing both a description of a judicial behavior and the observer's experience or reaction to the behavior), comments that could be improved (those that are codable but that are insufficiently elaborated to be included in the synthesis), and uninterpretable comments (those that cannot be coded as written).

Diversity of observers. Volunteers for courtroom observation are not necessarily representative of courtroom participants. The current cadre of observers are predominantly Caucasian, middle-aged or older, middle class, retired, and well educated. Because research suggests that the core principles of procedural fairness are universal, an unrepresentative cadre of volunteers may not adversely affect the validity of the observations. While research has shown that certain personalities are more sensitive to procedural fairness than others,³⁶ it is not known whether in fact an unrepresentative cadre of volunteers does experience certain judicial behaviors in the same way as courtroom participants. If the efforts of JPEC to recruit a more diverse cadre of volunteers are successful, then further research on the accumulating body of content analyses may shed light on this issue.

A second issue concerns the gender of judges and observers. Research has shown that female judges score consistently and significantly lower than their male counterparts in JPE surveys.³⁷ For example, female judges are systematically evaluated as weak and indecisive when observers perceive that they fail

33. See generally JANICE M. MORSE & LINDA NIEHAUS, *MIXED METHOD DESIGN: PRINCIPLES AND PROCEDURES* (2009).

34. Courtroom observation is undoubtedly daunting to many judges. JPEC's efforts to conduct its pilot study and to solicit feedback from judges about the program's design, including the use of a professional research consultant and the structure and usefulness of the summary report, helped to alleviate concerns about JPEC's use of courtroom observation in JPE.

35. Utah Admin. Code R597-3-4. The standard became effective on November 23, 2011, and first affects Utah judges facing retention election in 2014.

36. Manfred Schmitt & Martin Dörfel, *Procedural Injustice at Work, Justice Sensitivity, Job Satisfaction and Psychosomatic Well-Being* 29 *EUR. J. SOC. PSYCHOL.* (1999).

37. Wood, *et al.*, *supra* n. 9.

to assert control, but as unduly punitive when they assert authority. In one study, male judges were found to have a lower baseline expectation for courtesy, and so were judged as more courteous than female judges. Other research has shown that ingratiation between males and females is viewed with suspicion.³⁸ Where ingratiation is sincere and nonmanipulative, such as appropriate apology or praise, this does not threaten procedural fairness, but if a judge and the recipient of ingratiation are of different genders, this may have negative procedural-fairness effects. JPEC attempts to minimize potential gender effects by allocating both male and female observers to each courtroom wherever feasible.³⁹

Justice-court judge JPEs. Although full-time justice-court judges are currently evaluated in the same way as all other non-appellate judges in Utah, JPEC has found that these evaluations cannot be used with part-time justice-court judges. Part-time justice courts, with their low volume of cases, rural locations, part-time hours, high levels of self-represented litigants, and few court staff, make valid surveys results impossible to obtain. Courtroom observation will, however, remain a critical component of the justice court JPE, both because it is not dependent on large sample sizes and because of its particular value to part-time justice-court judges. Most litigants in justice courts do not have an attorney to represent them. Consequently, the direct interactions between judges and litigants are an especially important feature of a justice-court experience. The detailed procedural-fairness feedback contained in the courtroom-observation summary report could be some of the most valuable self-improvement information the JPE can provide to a justice-court judge.

Research on Accumulating Content Analyses of Courtroom Observations

As the number of content analyses that have been made increases, JPEC may decide to compare various groups of content analyses (e.g., by type of court, rural or urban location, years of experience, and gender of judge, etc.), and to explore themes, trends, and other patterns and commonalities within and among the various groups. This may include comparing and contrasting the quantitative survey scores for each judge with the qualitative content analyses of observation data, both to validate the evaluation process and to gain further understanding of judicial performance and evaluation.

Development of Procedural Fairness Based Judicial Training

The accumulating findings of the courtroom-observation program could be used to develop more judicial training about procedural fairness. The detailed descriptions of observed courtroom situations coupled with the observers' accounts of how they experienced those situations could be used in judicial education. Further, the systematic collection of courtroom practices that link specific judicial behaviors with procedural-fairness reactions may help to build a body of evidence about procedural fairness that judges find persuasive and compelling.

CONCLUSION

Utah's JPEC has taken advantage of the discretion afforded by the Utah Legislature to make administrative rules regarding the mandated courtroom-observation component of the Utah JPE program. Specifically, JPEC has focused courtroom observations on the procedural-fairness aspects of judicial behavior; established a minimum performance standard for procedural fairness; used qualitative methods to capture the rich, context-specific information available in the observers' narratives; and retained a consultant experienced in qualitative-evaluation research to systematically analyze and produce syntheses of the narrative data. JPEC hopes that its work will provide voters in retention elections with valued procedural-fairness information; provide members of the public, as volunteer observers, with a voice in JPE; provide judges with candid feedback on procedural-fairness behaviors for self-improvement; draw attention to issues of procedural fairness in the judiciary as a whole; and provide JPEC with a wider range of information for making retention recommendations. JPEC plans to continue its efforts to improve the courtroom-observation program, and looks forward to dialogue with other states about this important aspect of JPE.



Nicholas H. Woolf is the qualitative-research consultant to the Utah Judicial Performance Evaluation Commission; in 2011, he completed content analyses of courtroom-observation data for 72 Utah judges. Woolf has a Ph.D. in education from the University of Iowa and has served as qualitative research and evaluation consultant to numerous national and international projects in a variety of fields. He is a Visiting Academic Fellow at Henley Business School, part of the University of Reading in the United Kingdom. As an independent educator and trainer, he has conducted more than 280 workshops in qualitative-data analysis using the ATLAS.ti software program. More information about his work is available at www.learnatlas.com, and he may be contacted at woolf@pobox.com.



Jennifer MJ Yim, MPA, is the vice-chairperson of the Utah Judicial Performance Evaluation Commission. She is the past chairperson of JPEC's Courtroom Observation Subcommittee and a past member of its Rules Subcommittee. Her work in the justice system has included service on Utah's Third Judicial District Nominating Commission (2003-2007), Utah's Youth Parole Authority (2001-2011), and as the director of the Utah Judicial Council's Task Force on Racial and Ethnic Fairness in the Legal System (1997-2001). She is currently an appointed member of the Magistrate Merit Selection Panel for the United States District Court for the District of Utah, a procedural-justice consultant to the Utah Division of Juvenile Justice Services, and a doctoral candidate in the Department of Political Science at the University of Utah.

38. Robert C. Liden & Terence R. Mitchell, *Ingratiation Behavior in Organizational Settings*, 13 *ACAD. MGMT. REV.* 572 (1988).

39. Research has reported similar effects for African-American judges. No African-American judges are in the current cycle of Utah JPEs.