

THE USE OF JUDICIAL PERFORMANCE EVALUATION TO ENHANCE JUDICIAL ACCOUNTABILITY, JUDICIAL INDEPENDENCE, AND PUBLIC TRUST

DAVID C. BRODY*

In the twenty-first century the American judiciary has been under repeated attacks from multiple sources.¹ Editorials published in the media have called for the removal of judges from the bench for decisions made in individual cases and specific issues.² Members of Congress and other politicians have condemned “activist” judges and proposed punitive sanctions against the judiciary for decisions made in high profile, emotionally charged cases.³ Special interest groups actively seek the removal of judges at all levels who make rulings contrary to the groups’ beliefs or interests.⁴ Dozens of websites are housed on the Internet that call for increased accountability of judges in general⁵ and for the removal of individual judges.⁶

In 2006, attacks on the judiciary included ballot measures supported by highly organized campaigns to amend several state constitutions.⁷

* Associate Professor, Criminal Justice Program, Washington State University.

1. Sandra Day O’Connor, Op-Ed., *The Threat to Judicial Independence*, WALL ST. J., Sept. 27, 2006, at A18; Ruth Walker, *O’Connor Assails ‘Pervasive Attacks’ on Judges and Judicial Independence*, HARV. L. TODAY, Feb. 2007, at 1.

2. Editorial, *Forget Judicial Independence, It’s Time for Judicial Accountability*, N. COUNTRY GAZETTE, Mar. 17, 2006, <http://www.northcountrygazette.org/articles/031706Accountability.html>; *The O’Reilly Factor* (Fox News Channel broadcast Sept. 7, 2007), available at <http://www.foxnews.com/story/0,2933,296027,00.html>.

3. Charles Babbington, *Senator Links Violence to “Political” Decisions*, WASH. POST, Apr. 5, 2005, at A4; Carl Hulse & David D. Kirkpatrick, *DeLay Says Federal Judiciary Has ‘Run Amok,’ Adding Congress Is Partly to Blame*, N.Y. TIMES, Apr. 8, 2005, § A, at 21; Francis J. Larkin, *The Variousness, Virulence, and Variety of Threats to Judicial Independence*, 36 JUDGES’ J., Winter 1997, at 4, 5-6; William C. Mann, *Frist: Judges Forcing Gay Marriage Debate*, AP ONLINE, June 20, 2004. For a list of judges who have been criticized and specific anti-judge activities, see American Judicature Society, *Judges Under Fire*, http://www.ajs.org/cji/cji_fire.asp (last visited Oct. 19, 2008).

4. American Judicature Society, *Judges Under Fire*, http://www.ajs.org/cji/cji_fire.asp (last visited Oct. 19, 2008).

5. See, e.g., Citizens for Judicial Accountability, <http://www.judicialaccountability.org/> (last visited Oct. 19, 2008); Center for Judicial Accountability, <http://www.judgewatch.org/> (last visited Oct. 19, 2008); PACleanSweep.com, <http://www.pacleansweep.com/> (last visited Oct. 19, 2008); Judicial Accountability Initiative Law 4 Judges, <http://www.jail4judges.org/> (last visited Oct. 19, 2008); Get Off The Bench, <http://www.getyourjusticelive.com/> (last visited Oct. 19, 2008); Victims-of-Law, <http://victimsoflaw.net/> (last visited Oct. 19, 2008); A Matter of Justice, <http://www.amatterofjustice.org/amoj/00index.cfm> (last visited Oct. 19, 2007).

6. See, e.g., Vote No Judge Munger, <http://www.votenojudgemunger.com/> (last visited Oct. 19, 2008); Do Not Vote For Judge Kristin Booth Glen!!!, <http://www.anusha.com/notoglen.htm> (last visited Oct. 19, 2008).

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

These measures included providing for criminal and civil sanctions against judges for erroneous rulings,⁸ subjecting judges to electoral recall,⁹ redrawing judicial election districts,¹⁰ and imposing retroactive term limits for appellate court judges.¹¹

When considered together, these attacks on the judiciary share a common element: the belief that judges are not sufficiently accountable for the decisions they make. This lack of accountability, it is posited, gives rise to an “activist judiciary” that is free to make law and public policy contrary to the public will without fear of any consequences.¹²

These attacks on the judiciary have met strong resistance from members of the judiciary and the organized bar.¹³ Judges and bar leaders generally accept that there needs to be some level of judicial accountability (particularly regarding state court judges).¹⁴ They emphasize, however, that any efforts to increase judicial accountability must not infringe upon the independence of judges to make decisions as they deem just and proper under the law and Constitution.¹⁵

8. South Dakota Amendment E, Judicial Accountability Initiative Law (commonly known as J.A.I.L. for Judges). Judicial Accountability Initiative Law 4 Judges, <http://www.jail4judges.org> (last visited Oct. 19, 2008).

9. Montana Constitutional Initiative 98 would have amended the Montana Constitution to provide for recall by petition of state court justices or judges for any reason. Montana Secretary of State, 2006 Ballot Issues, <http://sos.mt.gov/elb/archives/2006/CI/CI-98.asp> (last visited Oct. 19, 2008).

10. Oregon Measure 40 would have amended the Oregon Constitution to require all appellate court judges be elected by district rather than statewide. Measure 40, Explanatory Statement, http://www.oregonvotes.org/nov72006/military_vp/m40_es.pdf (last visited Oct. 19, 2008).

11. Colorado Amendment 40 would have applied retroactive term limits for appellate court judges in Colorado. Ballot Title Setting Board, Proposed Initiative 2005-2006 #90, <http://www.elections.colorado.gov/WWW/default/Initiatives/Title%20Board%20Filings/Results%2090.pdf> (last visited Oct. 19, 2008).

12. Edwin Meese III & Rhett DeHart, *The Imperial Judiciary—And What Congress Can Do About It*, 81 POL’Y REV. 54, 54 (Jan.-Feb. 1997).

13. *E.g.*, Colorado Bar Association, Vote No 40, <http://www.cobar.org/judges/> (last visited Oct. 19, 2008); Susan Marmaduke & Jim Mountain, *Parting Thoughts: Let’s Keep Politics Out of the Courts*, OREGON STATE BAR BULLETIN, Oct. 2005, <http://www.osbar.org/publications/bulletin/06oct/parting.html>. See also American Judicature Society, Welcome to the Center for Judicial Independence, <http://www.ajs.org/cji/default.asp> (last visited Oct. 19, 2008) (discussing unfair criticism of judges who issue unpopular decisions and advocating judicial independence); American Bar Ass’n, Q&A on J.A.I.L. for Judges, Dec. 2006, <http://www.abanet.org/media/youraba/200612/article06.html> (discussing State Bar of South Dakota efforts to defeat 2006 Amendment E ballot initiative in South Dakota).

14. Shirley S. Abrahamson, *Thorny Issues And Slippery Slopes: Perspectives On Judicial Independence*, 64 OHIO ST. L.J. 3, 4 (2003); Troy A. Eid, *Judicial Independence and Accountability: The Case Against Electing Judges*, 30 COLO. LAW. 71, 71 (2001); Sandra Day O’Conner & RonNell Anderson Jones, *Reflections on Arizona’s Judicial Selection Process*, 50 ARIZ. L. REV. 15, 23 (2008); Peter D. Webster, *Selection and Retention of Judges: Is There One “Best” Method?*, 23 FLA. ST. U. L. REV. 1 *passim* (1995); Michael A. Wolff, *Chief Justice Michael A. Wolff: 2006 State of the Judiciary Address*, 62 J. MO. B. 56, 57-58 (2006); James Andrew Wynn, Jr. & Eli Paul Mazur, *Judicial Diversity: Where Independence and Accountability Meet*, 67 ALB. L. REV. 775 *passim* (2004).

15. O’Connor & Jones, *supra* note 14, at 17; Penny J. White, *Judging Judges: Securing Judicial Independence by Use of Judicial Performance Evaluations*, 29 FORDHAM URB. L.J. 1053, 1056 (2002); Larkin, *supra* note 3, at 4.

2008] *THE USE OF JUDICIAL PERFORMANCE EVALUATION* 3

The debate over the proper limits on judicial accountability and judicial independence is not a recent development. The proper role for judges in the American system of government has been hotly debated for more than 200 years.¹⁶ Advocates of a strong, independent judiciary, including Alexander Hamilton, argued that the role of judges is to faithfully interpret the law and constitutions without consideration of outside factors such as politics or popular sentiment.¹⁷ To make fulfilling this role practical, it is argued that judicial independence must be protected by insulating the judiciary from popular election.¹⁸ On the other hand, proponents of a judiciary that is directly accountable to the public view judges as governmental policymakers operating in a democracy.¹⁹ As with other policymakers in government, it is argued, judges must be directly accountable to the public for their actions by means of periodic elections, regardless of the effect this may have on judicial independence.²⁰

This philosophical debate has a practical effect on how judges in each of the fifty states are selected and kept in office.²¹ Proponents of a highly accountable bench favor direct election of judges.²² Those who advocate a fiercely independent judiciary favor judges being appointed to their position.²³

In what has been perceived by many to be an appropriate compromise between those advocating for contested election or direct appointment, in 1940 Missouri adopted a “merit selection” system of selecting state supreme court and court of appeals judges.²⁴ Under the “Missouri Plan,” when a judicial vacancy arises a nominating commission provides a list of three candidates to the governor. The governor then appoints

16. White, *supra* note 15, at 1053-54; Paul D. Carrington, *Restoring Vitality to State and Local Politics by Correcting the Excessive Independence of the Supreme Court*, 50 ALA. L. REV. 397, 403-404 (1999); Charles Gardner Geyh & Emily Field Van Tassel, *The Independence of the Judicial Branch in the New Republic*, 74 CHI.-KENT L. REV. 31, 35-36 (1998); William H. Rehnquist, *Judicial Independence*, 38 U. RICH. L. REV. 579, 582-83 (2004).

17. Geyh & Van Tassel, *supra* note 16, at 48.

18. *Id.* at 32.

19. PHILIP L. DUBOIS, FROM BALLOT TO BENCH 28-35 (1980); Jonathan L. Entin & Erik M. Jensen, *Taxation, Compensation, and Judicial Independence*, 56 CASE W. RES. L. REV. 965, 975 (2006) (citing BRUTUS, XV, in ESSAYS OF BRUTUS (Mar. 20, 1788), reprinted in 2 THE COMPLETE ANTI-FEDERALIST 437, 438 (Herbert J. Storing ed., 1981)).

20. Philip L. Dubois, *Accountability, Independence and the Selection of State Judges: The Role of Popular Judicial Elections*, 40 SW. L.J. 31, 38 (1986).

21. Webster, *supra* note 14, at 9.

22. David K. DeWolf, *Electing Judges Keeps Them Accountable*, SEATTLE POST-INTELLIGENCER, Nov. 3, 2006, at B7, available at http://seattlepi.nwsourc.com/opinion/290941_judgeelection03.html; Michael R. Dimino, *Judicial Elections Versus Merit Selection: The Futile Quest For a System Of Judicial “Merit” Selection*, 67 ALB. L. REV. 803, 815-18 (2004); Dubois, *supra* note 19, at 28.

23. Mark A. Behrens & Cary Silverman, *The Case for Adopting Appointive Judicial Selection Systems for State Court Judges*, 11 CORNELL J.L. & PUB. POL’Y 273, 304 (2002). See also Webster, *supra* note 14, at 13-14.

24. Webster, *supra* note 14, at 29-30.

one of the three candidates to fill the vacancy.²⁵ After a short period of time—generally a year—the new judge must stand before the voters in a retention election. If a majority of voters do not vote to have the judge remain on the bench, the judge is not retained and the process begins anew. If the judge is retained by the voters, he or she remains on the bench but is subject to periodic retention elections.²⁶

Proponents of the Missouri Plan argue that it takes politics out of the initial selection process but still provides for judicial accountability by requiring judges to stand for retention elections.²⁷ Opponents argue that since judges are rarely unsuccessful in surviving retention elections, any pretence of accountability derived from merit selection is a “sham.”²⁸ Central to this contention is the fact that voters have little information about the qualifications and performance of judges standing for retention.²⁹ Not coincidentally, voter participation rates in retention elections are nearly always significantly lower than rates in contested elections for other non-judicial elections held concurrently.³⁰

The lack of information about judges available for voters to use in retention elections and the effect this has on judicial accountability has been known to be present for a number of decades.³¹ In 1975, stemming from concern about voters lacking information to effectively vote in judicial retention elections,³² Alaska became the first state to take steps to address this issue by requiring the Alaska Judicial Council not only to evaluate judges appearing on a retention election ballot, but also to provide voters with the results of the evaluations along with recommendations as to whether each judge should be retained in office.³³ In 1976, the first judicial performance evaluation (JPE) took place. As of 2008, eight states provide their citizens with information obtained from official judicial performance evaluations of at least some of their judges standing for retention election for the express purpose of enabling voters to cast intelligent, meaningful votes.³⁴

25. *Id.* at 30.

26. *Id.*

27. Jay A. Daugherty, *The Missouri Non-Partisan Court Plan: A Dinosaur on the Edge of Extinction or a Survivor in a Changing Socio-Legal Environment?*, 62 MO. L. REV. 315, 318 (1997).

28. Dimino, *supra* note 22, at 806-07.

29. Webster *supra* note 14, at 34.

30. See William K. Hall & Larry T. Aspin, *What Twenty Years of Judicial Retention Elections Have Told Us*, 70 JUDICATURE 340, 342 (1987).

31. *Id.*

32. Telephone interview with Larry Cohn, Executive Director, Alaska Judicial Council (Aug. 8, 2008).

33. ALASKA STAT. §15.58.050 (2008).

34. The eight states include Alaska, Arizona, Colorado, Kansas, Missouri, New Mexico, Tennessee, and Utah. A number of other states operate judicial performance evaluation programs for self-improvement purposes. See INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., SHARED EXPECTATIONS: JUDICIAL ACCOUNTABILITY IN CONTEXT app. A (2006), available at <http://www.du.edu/legalinstitute/publications2006.html> [hereinafter SHARED EXPECTATIONS].

2008] *THE USE OF JUDICIAL PERFORMANCE EVALUATION* 5

Judicial performance evaluations are generally centered around responses to standardized, scaled surveys provided by individuals who have had direct dealings with a judge during an evaluation period.³⁵ The questionnaires ask these individuals, who may include attorneys, jurors, witnesses, court staff, and litigants, to rate the judge on behavior-based items related to process and demeanor, not outcomes.³⁶ The survey response data, along with other information such as court management data, recusal rates, courtroom observations, and disciplinary filings, are considered by a non-partisan commission made up of attorneys and lay persons.³⁷ After considering a judge's materials, the commission will come to a conclusion as to whether or not it believes that a judge should be retained in office. This final rating is presented to the public for use in retention election decisions.³⁸

The number of states, and in some cases counties, that use JPE programs continues to grow.³⁹ This growth is being fostered by the American Bar Association, which adopted guidelines for the establishment and operation of judicial performance evaluation programs in 2005.⁴⁰ The ability of JPE programs to provide voters with information they would otherwise lack is inarguably beneficial to conducting meaningful judicial elections. It must be noted, however, that support for judicial performance evaluation programs is premised on the assumption that, with the information JPE programs can provide, voters will be better equipped to hold judges appropriately accountable for their performance on the bench without interfering with judicial independence.⁴¹

35. Seth S. Andersen, *Judicial Retention Evaluation Programs*, 34 LOY. L.A. L. REV. 1375, 1380 (2001); David C. Brody, *The Relationship Between Judicial Performance Evaluations and Judicial Elections*, 87 JUDICATURE 168, 170 (2004); Kevin M. Esterling, *Judicial Accountability the Right Way*, 82 JUDICATURE 206, 209 (1999); Rebecca Love Kourlis & Jordan M. Singer, *Using Judicial Performance Evaluations to Promote Judicial Accountability*, 90 JUDICATURE 200, 201 (2007).

36. Andersen, *supra* note 35, at 1380; Brody, *supra* note 35, at 170; Esterling, *supra* note 35, at 209; Kourlis & Singer, *supra* note 35, at 201.

37. See SHARED EXPECTATIONS, *supra* note 34, at app. A.

38. *Id.*

39. Kansas recently conducted its first statewide JPE program for use in the 2008 retention elections. See The Kansas Judicial Report Card, <http://kansasjudicialperformance.org/> (last visited Oct. 19, 2008). Similarly, in 2008, Pierce County, Washington provided voters with the results of a JPE program that had attorneys and jurors evaluating judges before whom they appeared. Adam Lynn, *Judging the Judges: Bar Association Survey Rates Pierce County Judiciary*, THE NEWS TRIB. (Tacoma, Wash.), June 1, 2008, at A1, available at <http://www.thenewstribune.com/news/local/story/376713.html>. In North Carolina, the North Carolina Bar Association is conducting two pilot studies implementing a JPE program for the state's trial court judges, with plans to launch an actual evaluation program for use by voters in judicial elections for the 2010 elections. Interview with Allan Head, Executive Director, North Carolina Bar Association, in Cary, N.C. (June 23, 2008). Two other states, Minnesota and Indiana have committees examining the feasibility of conducting JPE programs in the near future. Interview with Jordan Singer, Director of Research, Institute for the Advancement of the American Legal System, in Denver, Colo. (August 5, 2008).

40. ABA, GUIDELINES FOR THE EVAL. OF JUDICIAL PERFORM. WITH COMMENTARY app. I (2005), http://www.abanet.org/jd/lawyersconf/pdf/jpec_final_commentary.pdf.

41. Kourlis, *supra* note 7, at 203; White, *supra* note 15, at 1061; Roger Handberg, *Judicial Accountability and Independence: Balancing Incompatibles?*, 49 U. MIAMI L. REV. 127, 135 (1994).

This assumption, however, is still just an assumption. While logical, it assumes that the presence of several necessary endogenous items. For a JPE program to fulfill its aim of increasing judicial accountability, the information provided to the public must be received by potential voters, regarded as trustworthy and reliable by the electorate, and used by voters in deciding how to vote in judicial elections. Additionally, to be effective, a JPE program needs to be accepted and trusted by the judges subject to evaluation.⁴² If the judiciary does not have confidence in the process or believes that it unduly interferes with judicial independence, then it may undermine the program's credibility with the public and present political or administrative obstacles that retard the short- and long-term operation of the program. Deficiencies in these areas will limit the program's ability to foster both judicial accountability and judicial independence.

To date, systematic assessments of whether JPE programs have any direct or indirect effects on judicial accountability, independence, and related items have been lacking. This article will attempt to go beyond the theoretical and consider how states can and should assess the effectiveness of judicial performance evaluation programs. After examining the concepts of judicial accountability and independence, and their relation to state court judicial elections, the article discusses the evolution and operation of judicial performance evaluations and the impact they may have on judicial accountability. The article then examines the specific impacts JPE programs may have on judicial accountability and sets forth criteria and means for examining if such impacts exist. The article concludes by examining the importance of evaluating both the methods used in conducting judicial performance evaluations and the impact the evaluations have on judicial accountability.

I. JUDICIAL INDEPENDENCE AND JUDICIAL ACCOUNTABILITY

The presence of judicial independence and judicial accountability are critical to functioning of America's constitutional democracy.⁴³ Absent either, the American constitutional system of checks and balances would likely collapse. The following section briefly discusses the definitions of judicial independence and accountability. This will lay an appropriate groundwork for my contention that rather than being contradictory values that are inherently at odds, they are in fact mutually supportive.

42. Brody, *supra* note 35, at 177.

43. See Bruce Fein & Burt Neuborne, *Why Should We Care About Independent and Accountable Judges?*, 84 JUDICATURE 58, 59 (2000); Jeffrey D. Jackson, *The Selection of Judges in Kansas: A Comparison of Systems*, 69-JAN J. KAN. B.A. 32, 35-37 (2000); Ronli Sifers, *Weighing Judicial Independence Against Judicial Accountability: Do the Scales of the International Criminal Court Balance?*, 8 CHI.-KENT J. INT'L & COMP. L. 88, 96 (2008); Roger K. Warren, *Judicial Accountability, Fairness, and Independence*, 42 CT. REV. 4, 7 (2005).

A. *Judicial Independence*

Judicial independence has been seen as a critical aspect of the American judicial system since before the formation of the nation.⁴⁴ During the colonial era, judges in the American colonies were appointed by the King of England and served at his will.⁴⁵ The control the King had over the judiciary and the direct control he had over the decisions handed down by colonial judges were sources of great concern and frustration for the colonists.⁴⁶ The injustices brought about by this system of “telephone justice”⁴⁷ angered the colonists to the extent that the lack of judicial independence of colonial judges was one of the enumerated grievances raised in the Declaration of Independence.⁴⁸ Following the colonial experience, it should not be surprising that the tripartite government premised on separation of powers and checks and balances would take great pains to provide for a judiciary that was coequal with and independent of the legislative and executive branches.

Legal scholars and social scientists have examined and defined judicial independence in myriad ways.⁴⁹ Conceptually it can be defined as having judges that

[A]re free to decide cases fairly and impartially, relying only on the facts and the law. It means that judges are protected from political pressure, legislative pressure, special interest pressure, media pressure, public pressure, financial pressure, and even personal pressure.⁵⁰

Judicial independence should be seen as a means to an end.⁵¹ Beyond its own value, it paves the way for a judiciary that treats people

44. THE FEDERALIST NO. 78, at 468 (Alexander Hamilton) (Clinton Rossiter ed., 1961). See also Berkeley N. Riggs & Tamera D. Westerberg, *Judicial Independence: An Historical Perspective the Independence of Judges Is . . . Requisite to Guard the Constitution and the Rights of the Individuals . . .*, 74 DENV. U. L. REV. 337, 351 (1997); Joseph H. Smith, *An Independent Judiciary: The Colonial Background*, 124 U. PA. L. REV. 1104, 1105 (1976).

45. Kelly J. Varsho, *In the Global Market for Justice: Who is Paying the Highest Price for Judicial Independence?*, 27 N. ILL. U. L. REV. 445, 447 (2007).

46. *Id.*

47. *Id.*

48. “[King George] has made Judges dependent on his Will alone, for the tenure of their offices, and the amount and payment of their salaries.” THE DECLARATION OF INDEPENDENCE para. 11 (U.S. 1776).

49. See Charles Gardner Geyh, *The State of the Onion: Peeling Back the Layers of America's Ambivalence Toward Judicial Independence*, 82 IND. L.J. 1215, 1223 (2007) (defining judicial independence as a tradition gained over time); G. Alan Tarr, Dep’t of Political Sci., Center for State Constitutional Studies, Rutgers Univ., Presentation of Paper at the Annual Meeting of the Midwest Political Science Association: Rethinking Judicial Independence and Judicial Accountability 3-4 (Apr. 12-15, 2007), <http://www.allacademic.com/one/mpsa/mpsa07/index.php> (search “Rethinking Judicial Independence”).

50. David J. Beck, *Judicial Independence: Woe to the Generation that Judges the Judges*, 71 TEX. B.J. 572, 572 (2008).

51. Kathryn Reed Edge, *Judicial Independence Judicial Accountability: A Difficult Balance*, TENN. B.J. May-June 1998, at 14; Warren, *supra* note 43, at 5; Frances Kahn Zemans, *The Accountable Judge: Guardian Of Judicial Independence*, 72 S. CAL. L. REV. 625, 632 (1999).

equally under the law.⁵² While complaints about activist judges and the need to rein in their independence are frequently raised, it is generally accepted that some degree of judicial independence is a necessary component of American government. This acceptable level of independence centers on the desire for judges to follow the law, to hold the other branches of government accountable to the Constitution, to treat parties that come before the bench equally and with impartiality, and to be free to apply the rule of law without fear of repercussion from the other branches of government.⁵³

B. *Judicial Accountability*

Just as it is generally accepted that judges must have some degree of independence to decide matters based on their interpretation of the law and facts presented before them, it is also widely accepted that judges should be held to some manner of accountability.⁵⁴ An accountable judiciary is important for a number of reasons. An unaccountable court system has a difficult time being perceived as legitimate and maintaining the trust and respect of the citizenry.⁵⁵ Absent public trust, the judiciary, which is dependent on other branches of government to enforce its orders and fund its operation, is greatly imperiled. Moreover, declines in trust and respect for a court system may ultimately lead to the loss of legitimacy in the eyes of the public and other branches of government. Such occurrences make repeated attempts to curtail judicial independence and authority, such as was seen in the 2006 ballot propositions, much more likely.

While judicial accountability is appropriate under the American system of government, the questions of who judges should be accountable to and what this accountability should be based upon depend largely on what is meant by judicial accountability. As noted by Charles Geyh, “[t]he peril of leaving judicial accountability ill-defined is that it can be co-opted and misused more easily.”⁵⁶ A meaningful method defining

52. Edge, *supra* note 51.

53. See Tarr, *supra* note 49, at 21 (acknowledging the general consensus that judges should not be punished “for following the law”); Zemans, *supra* note 51, at 634 (detailing the importance of impartiality for the judiciary).

54. Fein, *supra* note 43, at 59; Handberg, *supra* note 41 at 134; O’Connor & Jones, *supra* note 14, at 23; Wolff, *supra* note 14, at 57.

55. See Steven P. Croley, *The Majoritarian Difficulty: Elective Judiciaries and the Rule of Law*, 62 U. CHI. L. REV. 689, 710 (1995); Michael R. Dimino, *Pay No Attention to that Man Behind the Robe: Judicial Elections, the First Amendment, and Judges as Politicians*, 21 YALE L. & POL’Y REV. 301, 311-12 (2003); David E. Pozen, *The Irony of Judicial Elections*, 108 COLUM. L. REV. 265, 273 (2008); Mark S. Cady & Jess R. Phelps, *Preserving The Delicate Balance Between Judicial Accountability and Independence: Merit Selection in the Post-White World*, 17 CORNELL J. L. & PUB. POL’Y 343, 347 (2008).

56. Charles Gardner Geyh, *Rescuing Judicial Accountability from the Realm of Political Rhetoric*, 56 CASE W. RES. L. REV. 911, 912 (2006).

judicial accountability is to consider two types of performance criteria: *decisional accountability* and *behavioral accountability*.⁵⁷

1. Decisional Accountability

Decisional accountability involves holding judges answerable for their judicial decisions. In considering whether we should hold judges accountable for their decisions, there are several items that must be taken into account. First, was the decision contrary to established precedent? If the decision was contrary to established precedent, accountability can be maintained by appellate review. Were a judge to deliberately ignore precedent, state constitutions should provide for means of removing or disciplining the judge.

On the other hand, if a decision was reasonable and based on precedent, the desire to hold a judge accountable for an unpopular ruling is fraught with peril. It is the impulse to hold judges accountable for legally sound rulings by means of electoral challenges that demonstrates the danger of majoritarian rule premised on decisional accountability.⁵⁸

The dangers of holding judges directly accountable for making necessary but politically unpopular decisions have been noted throughout the nation's history by some of its most noted jurists.⁵⁹ In *Dennis v. United States*, Justice Felix Frankfurter wrote, "[c]ourts are not representative bodies. They are not designed to be a good reflex of a democratic society Their essential quality is detachment, founded on independence."⁶⁰ Similarly, Justice Black, in *Chambers v. Florida*, noted that "courts stand against any winds that blow as havens of refuge for those who might otherwise suffer because they are helpless, weak, outnumbered, or because they are non-conforming victims of prejudice and public excitement."⁶¹

The classic example of such a campaign involved the defeat of Tennessee Supreme Court Justice Penny White in the 1996 retention election.⁶² Justice White was appointed to the Tennessee Supreme Court in 1994. In 1996 Justice White concurred with an opinion which held that a

57. Wendell L. Griffen, *Judicial Accountability and Discipline*, 61 *LAW & CONTEMP. PROBS.* 75, 76 (1998). A third type of judicial accountability, institutional accountability, involves the accountability of the judiciary as a branch of government. Institutional accountability concerns items such as being responsible for maintaining a reasonable budget, operating in an efficient manner, and hearing and deciding matters in an acceptably speedy manner given its resources. See Geyh, *supra* note 56, at 917-19.

58. See Croley, *supra* note 56, at 725-29.

59. Judge Peter Webster compiles several of these statements. See Peter D. Webster, *Who Needs an Independent Judiciary?*, *FLA. B.J.*, Feb. 2004, at 24, 25-28.

60. 341 U.S. 494, 525 (1951) (Frankfurter, J., concurring).

61. 309 U.S. 227, 241 (1940).

62. See Jeffrey D. Jackson, *Beyond Quality: First Principles in Judicial Selection and Their Application to a Commission-Based Selection System*, 34 *FORDHAM URB. L.J.* 125, 134 (2007); Traciell V. Reid, *The Politicization of Retention Elections Lessons from the Defeat of Justices Lanhier and White*, 83 *JUDICATURE* 68, 69-70 (1999).

murder committed during the commission of a rape was not automatically “heinous, atrocious, or cruel” and therefore may not serve as an aggravating circumstance warranting execution unless there was evidence that the act “involved torture or serious physical abuse beyond that necessary to produce death.”⁶³ A majority of the court, including Justice White, agreed that such evidence was not present in the case, vacated the death sentence, and remanded the case for resentencing.⁶⁴

In 1996 Justice White was the only member of the Tennessee Supreme Court scheduled to stand for retention election. In response to *Odom*, a coalition of conservative organizations, victim’s rights groups, law enforcement agencies, and others organized to mount a campaign to defeat Justice White. In the end, Justice White was defeated, garnering only 45% of the vote in favor of her retention. The defeat of Justice White has been widely viewed as a direct attack on judicial independence based on the improper use of decisional accountability.⁶⁵

2. Behavioral Accountability

As opposed to decisional accountability, behavioral accountability involves holding individual judges answerable for their conduct on the bench.⁶⁶ As judges are the human element of the justice system, conduct which reflects badly on the integrity and impartiality of the justice system is likely to decrease public trust in the judiciary. Explicit statements or acts of bias and partiality, *ex parte* communications, rudeness, and a lack of respect for parties or counsel, are examples of actions for which a judge may be held accountable.⁶⁷ Acts related to behavioral accountability are universally accepted as being appropriate components of judicial accountability and do not restrict appropriate aspects of judicial independence. Aspects of a judge’s performance on the bench are appropriate foci for voters or reappointing authorities to consider in whether a judge should receive another term on the bench. While appropriate, such information is difficult for voters to obtain and examine at a broad enough level to promote judicial accountability.

C. *The Potential of Co-Occurring Judicial Independence and Judicial Accountability*

Judicial independence and judicial accountability have frequently been viewed as competing forces which need to be counterbalanced against one another.⁶⁸ On one hand, judges who possess excessive inde-

63. *State v. Odom*, 928 S.W.2d 18, 24, 26 (Tenn. 1996).

64. *Id.* at 32-33.

65. *See Zemans*, *supra* note 51, at 648-49.

66. Geyh, *supra* note 56 at 919.

67. White, *supra* note 15 at 1064-1070.

68. Michael R. Dimino, *supra* note 22 at 803; Tillman J. Finley, Note, *Judicial Selection in Alaska: Justifications and Proposed Courses of Reform*, 20 ALASKA L. REV. 49, 49 (2003); David Schultz, *Minnesota Republican Party v. White and the Future of State Judicial Selection*, 69 ALB. L.

2008] *THE USE OF JUDICIAL PERFORMANCE EVALUATION* 11

pendence are free to act without concern for being held accountable for overstepping their authority, employing expansive interpretations of the Constitution, or simply ignoring or nullifying the law when they see fit.⁶⁹ On the other hand, if a judiciary is overly accountable to the public, then judges will be fearful of making correct legal decisions that are contrary to the public will.⁷⁰

In reality, judicial independence and accountability are not at odds with each other but are rather “different sides of the same coin,”⁷¹ “yin” to each other’s “yang,”⁷² and “both desirable.”⁷³ In fact, rather than being seen as counterbalances to each other, it is possible for them to be mutually beneficial.

For courts to operate effectively, citizens must have trust and confidence in judges to decide matters in a fair and impartial manner.⁷⁴ Courts have no means to enforce their own orders. The public’s use of the court system to resolve disputes and its compliance with judicial orders is dependent on support for courts as legitimate institutions.⁷⁵ This legitimacy hinges to a large extent on public trust in the judiciary as an institution and judges as ministers of justice. A trusted judiciary will be respected and viewed as legitimate. A judiciary that is not trusted may have its legitimacy, authority, and eventually orders questioned by the citizenry or by the other branches of government.⁷⁶

An important component of public trust in governmental institutions is the citizenry’s ability to hold its officers accountable.⁷⁷ A governmental institution that is viewed as unaccountable is likely to lack public trust. On the other hand, as public accountability increases, the trust afforded the institution is likely to increase as well.⁷⁸

Beyond increasing perceived legitimacy, the public’s ability to effectively hold judges accountable has the potential to generate what Tom Tyler calls a “reservoir of loyalty”⁷⁹ upon which the judiciary can draw. This reservoir grows as the public’s trust in the courts and judiciary operate effectively in a transparent manner. After time, this reservoir can

REV. 985, 987 (2006); Anthony Champagne, *The Selection and Retention of Judges in Texas*, 40 SW. L.J. 53, 111-17 (1986).

69. See *supra* notes 2-6.

70. See Croley, *supra* note 55 at 908.

71. Stephen B. Burbank, *The Architecture of Judicial Independence*, 72 S. CAL. L. REV. 315, 339 (1999).

72. Geyh, *supra* note 56 at 916.

73. SHARED EXPECTATIONS, *supra* note 34 at 6.

74. George W. Dougherty, Stefanie A. Lindquist, & Mark D. Bradbury, *Evaluating Performance in State Judicial Institutions: Trust and Confidence in the Georgia Judiciary*, 38 ST. & LOC. GOV’T REV. 176, 176 (2006).

75. *Id.* at 177.

76. *Id.*

77. *Id.*

78. TOM R. TYLER, *WHY PEOPLE OBEY THE LAW* 26 (1990).

79. *Id.*

be drawn upon by judges to foster the acceptance and understanding of politically unpopular decisions that courts are required to make. Such public acceptance in turn serves to increase judicial independence and the ability to make difficult decisions without concern for political ramifications.

While this mutually beneficial series of events is theoretically possible, it first requires a foundation of judicial independence accompanied by appropriate and effective means of judicial accountability. The basis upon which voters hold judges accountable must not infringe upon their ability to decide matters based on the law. While judicial elections provide the vehicle to provide judicial accountability to the public and to provide true accountability that will increase independence, voters need to be given sufficient information with which to make electoral decisions. This dynamic, and the role judicial performance evaluation programs have in it, will be discussed in later sections.

II. JUDICIAL ELECTIONS AND ACCOUNTABILITY

In thirty-nine states, judges are theoretically held accountable⁸⁰ by the public through either contested elections or retention elections.⁸¹ As noted previously, whichever mode of election a state uses to choose or retain its judges, elections allow the public to hold judges accountable by use of democratic means. That being said, simply having an election does not mean it will provide meaningful levels of judicial accountability.

For judicial elections to effectively hold judges accountable, several things are required. First, it is important that the electorate participate in judicial elections at a reasonably high level.⁸² Acceptable turnout levels are needed for two reasons: 1) to have the election considered legitimate and 2) to serve as a tool which judges must respect. Elections with dismal levels of participation weaken the perception that they serve as a genuine means to hold government actors accountable for their job performance.⁸³ As Michael Dimino, a strong advocate of judicial elections, put it, "If judges are to receive the benefit of legitimacy that comes from

80. While beyond the scope of this article, it is worth noting that states that use partisan and non-partisan elections as a means of selecting and retaining judges rarely have contested elections. For example, in 2008 in Washington a state that selects judges via non-partisan elections, 84% of the judges subject to reelection ran unopposed, removing any semblance of public accountability. This level of competition is typical in non-partisan election states.

81. Rachel Caufield, *In the Wake of White: How States are Responding to Republican Party of Minnesota v. White and How Judicial Elections are Changing*, 38 AKRON L. REV. 625, 629 (2005). Throughout this article when the generic term "judicial elections" is used, it refers to both contested judicial elections and retention elections.

82. David Adamany & Phillip Dubois, *Electing State Judges*, 1976 WIS. L. REV. 731, 736 (1976).

83. STEVEN E. SCHIER, *YOU CALL THIS AN ELECTION?: AMERICA'S PECULIAR DEMOCRACY* 74 (2003).

2008] *THE USE OF JUDICIAL PERFORMANCE EVALUATION* 13

having periodic elections, it seems that elections should encourage participation by as many eligible voters as possible.”⁸⁴

Low turnout in judicial elections also has the potential to affect the way judges do their jobs. Should judges know that only a small percentage of potential voters are going to participate in a judicial election, the number of people a judge is actually accountable to decreases. Moreover, if the demographic, social, or political characteristics of voters likely to vote in low turnout elections are known, a judge can adjust his or her behavior to satisfy the preference of those individuals.⁸⁵ In this respect low levels of turnout may serve to negate any increases in perceived accountability and legitimacy of the judiciary obtained through judicial elections.

A key aspect of increasing turnout and the effectiveness of judicial elections is providing voters with sufficient amounts of relevant information to use in casting their votes.⁸⁶ Research has found that low levels of turnout in judicial elections are due largely to the fact that voters know little, if anything, about the names appearing on their ballots.⁸⁷ When voters have no information about the candidates on an election ballot, their votes are frequently based on heuristics or voting cues.⁸⁸ This leads to election outcomes that are based on name recognition, gender, ethnic preferences, ballot position, or pure luck.⁸⁹

The lack of information about candidates makes it fair to conclude that “judicial accountability through the election process is minimal.”⁹⁰ This is so because elections are, for the most part, contests in which the public has very little information about the candidates and relatively few voters choose to participate. While advocates of judicial elections may argue that elections do provide for judicial accountability, there are several undeniable facts that refute this position.

It is beyond dispute that judicial elections suffer from high levels of voter “falloff.” Falloff is defined as “[t]he difference between how many people go to the polls and how many people actually vote on a specific

84. Dimino, *supra* note 55 at 374.

85. Research has consistently shown that in low turnout elections, those who do vote are more likely to better educated and significantly wealthier than non-voters. Georg Lutz & Michael Marsh, *Introduction: Consequences of Low Turnout*, 26 *ELECTORAL STUDIES* 539, 543 (2007).

86. Martin P. Wattenberg, Ian McAllister, and Anthony Salvanto, *How Voting is Like Taking the SAT Test: An Analysis of American Voter Rolloff*, 28 *AM. POL. Q.* 234, 247 (2000).

87. Matthew J. O'Hara, Student Note and Comment, *Restriction of Judicial Election Candidates' Free Speech Rights After Buckley: A Compelling Constitutional Limitation?*, 70 *CHI.-KENT L. REV.* 197, 209-210 (1994); Charles H. Sheldon & Nicholas P. Lovrich, Jr., *Knowledge and Judicial Voting: The Oregon and Washington Experience*, 67 *JUDICATURE* 235, 239 (1983).

88. Sheldon & Lovrich, Jr., *supra* note 87.

89. *Id.* at 235; see also Larry Aspin et al., *Thirty Years of Judicial Retention Elections: An Update*, 37 *SOC. SCI. J.* 1, 3 (2000); Marie Hojnacki & Lawrence Baum, *Choosing Judicial Candidates: How Voters Explain Their Decisions*, 75 *JUDICATURE* 300, 308-09 (1992).

90. Handberg, *supra* note 41 at 132.

[contest].”⁹¹ This occurs because in most elections there are a number of matters for which voters are asked to cast votes. These contests may range from president and governor, to school board and water commissioner. Judicial elections are usually among the lower tier of contests appearing on a ballot. Normally, some people who vote in top level contests do not vote in other races appearing lower down on the ballot, including judicial elections.

Falloff, also referred to as “dropoff”, “rolloff” and “ballot fatigue” in the legal and academic literature,⁹² is not a new development.⁹³ Lawrence Baum reported that in judicial elections between 1980 and 1995 roughly 25% of voters who cast ballots in an election did not vote in state supreme court elections.⁹⁴ The falloff in these elections was over 20% regardless of whether they involved contested or retention elections.⁹⁵

While a number of factors can affect falloff levels, their primary cause is a lack of information possessed by voters about the given contest.⁹⁶ This phenomenon has been compared to how college applicants approach the SAT⁹⁷ exam in which points are deducted for incorrect answers, but there is no penalty for skipped questions.⁹⁸ As such, it is prudent for test takers to skip over questions for which they lack the necessary knowledge to answer correctly. This is a strategy frequently taken by voters in judicial elections; absent any information about the judge or judges appearing on the ballot, voters skip over the contest and move onto the next.⁹⁹

Of course, not all voters who lack information about judicial candidates refrain from voting. A large number of people who lack any meaningful information about the judge standing for retention or the candidates vying for a contested position nonetheless vote in the contests. Despite the increased level of spending on judicial campaigns and decreased restrictions on judicial campaign speech, the amount of information used by voters in judicial elections has remained low.¹⁰⁰ In fact, having judicial campaigns look more like other election campaigns may

91. Mathew Manweller, *Examining Decreasing Rates of Voter Falloff in California and Oregon*, 36 ST. & LOC. GOV'T. REV. 59, 59 (2004).

92. *Id.* at 65 n.2; see also MARK LAWRENCE KORNBLUH, *WHY AMERICA STOPPED VOTING: THE DECLINE OF PARTICIPATORY DEMOCRACY AND THE EMERGENCE OF MODERN AMERICAN POLITICS* 89-105 (2000).

93. ANTHONY DOWNS, *AN ECONOMIC THEORY OF DEMOCRACY* 260-67 (1957).

94. Lawrence Baum, *Judicial Elections and Judicial Independence: The Voter's Perspective*, 64 OHIO ST. L.J. 13, 19 (2003).

95. *Id.* at 19-20 (finding that partisan elections averaged 22% falloff, while nonpartisan and retention elections averaged 29% and 28% falloff, respectively).

96. Wattenberg et al., *supra* note 86, at 236.

97. The SAT Reasoning Test, formerly known as the Scholastic Aptitude Test.

98. Wattenberg et al., *supra* note 86, at 236.

99. *Id.* at 236-37.

100. Steven Zeidman, *To Elect or Not to Elect: A Case Study of Judicial Selection in New York City 1977-2002*, 37 U. MICH. J.L. REFORM 791, 819-21 (2004).

have the effect of decreasing respect for the court as an institution¹⁰¹ and voter participation.

These facts create a bit of a conundrum. While the public likes having the ability to hold judges electorally accountable and wants more information to use in judicial elections, at the same time, it wants judges to be above politics, sufficiently independent to base rulings on the law and not public opinion, and insulated from special interest groups.¹⁰² In the following section, I discuss how judicial performance evaluations of judges standing for retention or reelection may increase the effectiveness of judicial elections in holding judges accountable.

III. JUDICIAL PERFORMANCE EVALUATIONS

As discussed above, judicial performance evaluations were first used in the 1970s to provide information for voters to use in judicial retention elections. Since that time nearly twenty states have implemented official JPE systems, with eight doing so to provide information to voters.¹⁰³ Typically, JPE programs are based on evaluations of how well a judge demonstrates a number of qualities expected of an excellent jurist submitted by individuals who have experience appearing before the judge. These items generally fall into the categories of legal ability, integrity, communication, judicial temperament, and administrative ability.¹⁰⁴ The keys to JPE systems are that they 1) involve information only from individuals who have first-hand knowledge, through observation, of a judge's performance and 2) expand the sources of information beyond attorneys to include lay persons, jurors, witnesses, and court staff who have served, testified, or worked in a judge's court and had the ability to personally observe the judge's performance. This information is then considered by a commission which makes a retention recommendation to the public for use in considering how to vote in judicial elections.¹⁰⁵

The information provided to voters by JPE programs has the potential to provide a missing ingredient for judicial elections to appropriately and effectively facilitate judicial accountability. Proponents of an elected judiciary and merit selection systems both agree that the lack of information available to voters is a serious problem associated with judicial elections.¹⁰⁶ Beyond affecting voter participation, the lack of infor-

101. Mark A. Behrens & Cary Silverman, *The Case for Adopting Appointive Judicial Selection Systems for State Court Judges*, 11 CORNELL J.L. & PUB. POL'Y 273, 282-85, 287 (2002).

102. See Pozen, *supra* note 55, at 271-72.

103. See SHARED EXPECTATIONS, *supra* note 34, at app. a.

104. INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., TRANSPARENT COURTHOUSE: A BLUEPRINT FOR JUDICIAL PERFORM. EVAL. 13-14 (2006), <http://www.du.edu/legalinstitute/pubs/TransparentCourthouse.pdf>. The Institute for the Advancement of the American Legal System provides state-specific information about the criteria and methods used in the JPE programs operated in Arizona and Colorado. SHARED EXPECTATIONS, *supra* note 34, at app. b.

105. See SHARED EXPECTATIONS, *supra* note 34, at 10-11.

106. See Dimino, *supra* note 22, at 807; Dubois, *supra* note 19, at 32.

mation possessed by voters calls into question the legitimacy of electoral accountability.¹⁰⁷

As stated shortly after the implementation of Colorado's judicial performance evaluation program:

In order for citizens to maintain popular accountability of the judiciary, citizens must be involved in evaluating judicial performance They need to (1) gather information about judicial performance from the citizen's point of view, and (2) communicate their opinions to the judiciary. . . . [T]he main vehicle of judicial accountability is the [election or retention of judges]. Yet this . . . cannot serve its function if citizens do not have the interest to vote or the information necessary to make informed decisions.¹⁰⁸

The possession of reliable and relevant information by voters is an important ingredient to effective democracy.¹⁰⁹ For government to possess democratic legitimacy it is important that the electorate have the ability to "deliberate" before casting a vote.¹¹⁰ Moreover, it is important that the government makes the collection of such information easy for all citizens.¹¹¹ Absent such information, judicial elections provide little accountability and great peril.

The perceived ability of JPE programs to provide information to voters and to foster judicial accountability in retention election states has been widely applauded.¹¹² This positive reception is likely because a JPE program that can provide voters with standardized information about a judge's performance has the potential not only to increase judicial accountability, but to increase judicial independence as well.¹¹³

This position was succinctly summarized in 2006 by the former President of the Kansas Bar Association and current Chair of the Kansas Commission on Judicial Evaluation:

107. Alex B. Long, "Stop Me Before I Vote for this Judge Again": *Judicial Conduct Organizations, Judicial Accountability, and the Disciplining of Elected Judges*, 106 W. VA. L. REV. 1, 42 (2003).

108. White, *supra* note 15, at 1064 (quoting Anne Rankin Mahoney, *Citizen Evaluation of Judicial Performance: The Colorado Experience*, 72 JUDICATURE 210, 216 (1989)).

109. MICHAEL X. DELLI CAPRINI & SCOTT KEETER, WHAT AMERICANS KNOW ABOUT POLITICS AND WHY IT MATTERS 5-7 (1996); DENNIS F. THOMPSON, JUST ELECTIONS: CREATING A FAIR ELECTORAL PROCESS IN THE UNITED STATES 89 (2002); *cf.* Richard R. Lau & David P. Redlawsk, *Voting Correctly*, 91 AM. POL. SCI. REV. 585, 586 (1997) ("[I]f we are going to make judgments about the 'democratic' nature of different forms of government, we should do so at least initially on the basis of the quality or 'correctness' of the political decisions citizens make within that system of government rather than on the basis of the ways in which those decisions are reached.").

110. Loren A. King, *Deliberation, Legitimacy, and Multilateral Democracy*, 16 GOVERNANCE 23, 25-26 (2003) (defining deliberation as "a process of careful and informed reflection on facts and opinions, generally leading to a judgment on the matter at hand.").

111. *See id.* at 41-42.

112. A. John Pelander, *Judicial Performance Review in Arizona: Goals, Practical Effects and Concerns*, 30 ARIZ. ST. L.J. 643, 651-52 (1998).

113. White, *supra* note 15, at 1064, 1075-76.

Why do we need to evaluate our judges? One of the most frequent complaints about retention elections is that no one knows the record of a judge who is up for retention. This creates an information vacuum which is too inviting for groups with a philosophical axe to grind. We've seen recent instances in district court retention races in Douglas and Shawnee counties where narrow focus groups attacked judges running for retention by distorting the judge's performance record to suit the objectives of the attack group. Without unbiased information the voting public is faced with deciding judicial qualifications by sound bites.¹¹⁴

JPE programs are designed to fill this vacuum. Beyond providing information for use in holding judges accountable for their performance, JPE programs may foster judicial independence directly by providing information voters can use in judicial elections that is not issue-based, but rather on whether a judge does his or her job as one would expect from a judge.¹¹⁵ Reviews of judicial performance evaluation programs based primarily on anecdotal evidence have been generally, though not universally, positive. The assessments, however, are very general in nature and based largely on anecdotes. Given that JPE programs have been part of the judicial and electoral landscape for several decades and have had time to mature and evolve, it is time for independent evaluations of their impact at both the state and national levels. The balance of this article examines potential criteria for assessing the effectiveness of JPE programs in increasing judicial accountability, applies these criteria to select programs, and discusses the importance of such an assessment for enhancing the judicial independence and judicial accountability of state judges.

IV. ASSESSING JUDICIAL PERFORMANCE EVALUATION PROGRAMS

While the specific details surrounding the operation of individual judicial performance evaluation programs vary, a number of programs are charged with providing reliable, relevant information to the public for use in deciding how to vote in judicial elections.¹¹⁶ This objective is only a means to other ends. As discussed above, the broader objective behind these programs is to increase judicial accountability and independence by facilitating increased voter turnout and informed voting in judicial elections. Evaluating whether JPE programs are able to achieve these ends requires a mixed method that addresses several questions.¹¹⁷

114. Richard F. Hayse, *The Need for Judicial Performance Evaluation*, J. KAN. B. ASS'N, Apr. 2006, at 4.

115. SHARED EXPECTATIONS, *supra* note 34, at 7.

116. See SHARED EXPECTATIONS, *supra* note 34, at 13-16 and surrounding text.

117. See Jennifer C. Greene et al., *Toward a Conceptual Framework for Mixed-Method Evaluation Designs*, 11 EDUC. EVALUATION & POL'Y ANALYSIS 255 (1989) for an explanation of mixed-method evaluation design.

A. *Do Judges with Poor Evaluations Receive Fewer Positive Votes in Retention Elections?*

Logic would seem to indicate that if JPE programs are effective in holding judges accountable, then this would be evidenced by the behavior of voters in retention elections. More specifically, one would expect a positive relationship to be present between the rating a judge receives from a JPE survey or commission and the percentage of voters who vote that the judge should be retained in office. The manner in which JPE results are reported in the majority of states makes such an analysis difficult. Most states that conduct evaluations for use in retention elections do not provide a single numerical measure of a judge's performance. Instead, they either give a recommendation as to whether the judge should be retained without a numerical rating, or they provide ratings for multiple, independent criteria (such as legal ability, integrity, etc.). Absent ordinal rankings of judicial performance that are readily identifiable and sufficiently varied among judges, examining relationships between ratings and voter behavior is not practical.

There are two states—Alaska and Utah—that do provide voters with specific ratings for each judge standing for retention election that permit such an analysis.¹¹⁸ In an effort to consider the potential effectiveness of JPE programs in holding judges accountable, I examined the relationship between the evaluation ratings received from attorneys by judges standing for retention in the state of Alaska and the percentage of voters who voted that the judge should be retained in office.¹¹⁹ Figure 1 presents a scatter diagram of these items for the eighty-six judicial retention elections held in Alaska between 2000 and 2006. In the chart presented in Figure 1, each dot represents a judge standing for retention. The percentage of voters voting to retain the judge is measured along the vertical axis, while the overall rating received from attorneys completing evaluation questionnaires is measured along the horizontal axis.¹²⁰

118. In 2001, Susan Olson examined the impact evaluations had on the results of judicial retention elections involving judges in the Salt Lake City area over three election cycles. Susan M. Olson, *Voter Mobilization in Judicial Retention Elections: Performance Evaluations and Organized Opposition*, 22 JUST. SYS. J. 263, 267 (2001). While Olson found that the three judges who received poor evaluation ratings during that time period received fewer votes in favor of retention than their colleagues, she also reported that the active campaigns against one of the judges had a much greater impact on voters than did the judicial performance evaluations. *Id.* at 278.

119. Alaska election results obtained from the State of Alaska Division of Elections website. State of Alaska Division of Elections, <http://www.elections.alaska.gov> (last visited Oct. 19, 2008). Judicial evaluation ratings obtained from the Alaska Judicial Council website. Alaska Judicial Council, <http://www.ajc.state.ak.us> (last visited Oct. 19, 2008).

120. In addition to criteria specific behavior based questions, attorneys were asked to provide an overall performance rating for a judge ranging from 1 (unacceptable) to 5 (excellent). See, e.g., ALASKA JUDICIAL COUNCIL RECOMMENDATION 2 (2008), <http://www.ajc.state.ak.us/retent08/coats08.pdf>. The average of these ratings received by each judge is used in this analysis.

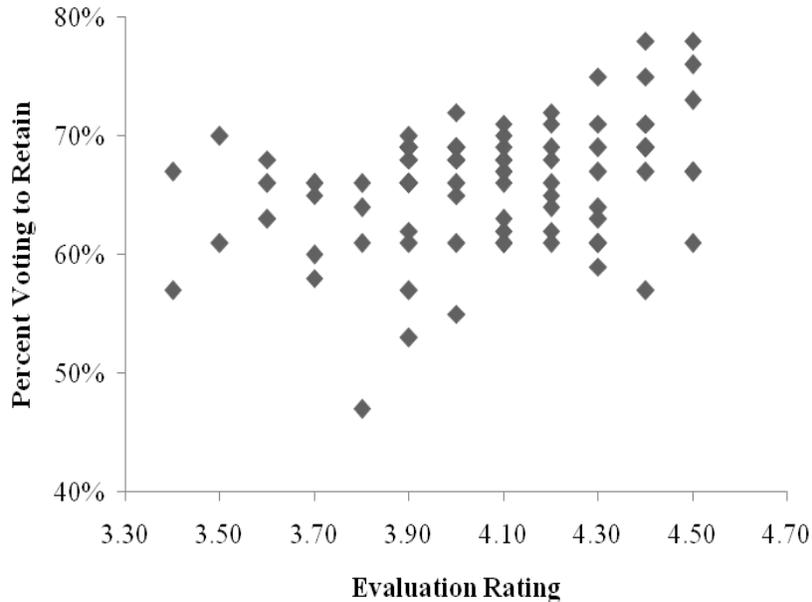


FIGURE 1 ALASKA JUDICIAL RETENTION ELECTIONS, 2000-2006

As can be readily seen in Figure 1, there is a significant linear relationship between the rating received as part of the JPE program and the percent voting that the judge should be retained (higher rating associated with increased yes votes received).¹²¹ While one cannot prove direct causation between evaluation ratings and election results, this simple analysis does suggest such an effect is present to some degree.¹²²

B. *Do Judges Who Receive Poor Evaluations Remain in Office?*

A fundamental question about the effectiveness of JPE programs in holding judges accountable is whether judges who receive poor evaluations are leaving the bench. While the question itself is straightforward, arriving at an accurate answer is not.

It is an undeniable fact that only a very small number of retention elections result in a judge being voted out of office.¹²³ Larry Aspin reports that between 1964 and 2006 only fifty-six of the 6,306 judges fac-

121. The statistically significant correlation between the two variables is .34.

122. Further multivariate analysis could be conducted controlling for possibly confounding variables including location of judgeship, gender, race, presidential election year, as well as many other state-specific factors.

123. Michael R. Dimino, *The Futile Quest for a System of Judicial "Merit" Selection*, 67 ALB. L. REV. 803, 807 (2004); Larry Aspin, *Judicial Retention Election Trends 1964-2006*, 90 JUDICATURE 208, 210 (2007).

ing judicial retention elections were defeated.¹²⁴ Moreover, there is no significant difference in retention election outcomes for states with judicial performance evaluation programs as opposed to states without.¹²⁵ Such statistics have frequently been used by opponents of merit selection as ammunition for their position that retention elections are ineffective in holding judges accountable.¹²⁶

At first glance, these statistics would seem to indicate judicial performance evaluation programs do little to actually hold judges accountable. Such a conclusion would be purely speculative. One must keep in mind that judicial accountability is premised on holding judges performing below acceptable standards accountable. It does not operate as a quota system or with a mandatory curve by which only a limited number of judges can be found to be performing well. As over 90% of judges evaluated across the country receive positive recommendations from their JPE commissions,¹²⁷ the fact that an equally high number of judges are retained in office in retention elections should not be surprising. One might even posit that it is appropriate. After all, if good attorneys are appointed to the bench, and they perform well as judges, voters should happily vote for retention.

From time to time judges do receive negative evaluations that recommend citizens to vote against retention. The election results in such instances have been mixed. In Alaska, of the two judges who have received non-retention recommendations since 1984, one was retained by the voters while one was voted off the bench.¹²⁸ In Colorado, of the six judges who have been removed by the voters in retention elections since 1988, five received “do not retain” recommendations.¹²⁹ Prior to 2008,¹³⁰ the Arizona Commission on Judicial Performance Review found only one judge did not meet judicial performance standards.¹³¹ The judge was subsequently retained by the voters.¹³² On the other hand, two Arizona judges who received positive recommendations from the commission were not retained by the electorate.¹³³

124. Aspin, *supra* note 123, at 210.

125. *Id.* at 213.

126. See John Andrews, *Judges Coddled by Sweetheart Process*, DENVER POST, Feb. 18, 2007, at E-3.

127. *Id.*

128. ALASKA JUDICIAL COUNCIL, SELECTING AND EVALUATING ALASKA'S JUDGES: 1984-2007, at 36 (2008), <http://www.ajc.state.ak.us/Reports/JudgeProfile08.pdf>.

129. E-mail from Jane Howell, Executive Director, Office of Judicial Performance Evaluation (Sept. 10, 2008) (on file with author).

130. In 2008, the Arizona Commission on Judicial Review recommended one Maricopa County Superior Court judge not be retained.

131. Mark I. Harrison, Sara S. Greene, Keith Swisher & Meghan H. Grabel, *On the Validity and Vitality of Arizona's Judicial Merit Selection System: Past, Present, and Future*, 34 FORDHAM URB. L.J. 239, 257 n.127 (2007).

132. *Id.*

133. *Id.* at 258.

2008] *THE USE OF JUDICIAL PERFORMANCE EVALUATION* 21

From the results noted above, it would appear JPE does little to remove poorly performing judges from the bench. In reality, however, it does. In states that use JPE in conjunction with retention elections, if a judge chooses not to seek another term in office, then his or her performance evaluation is not released to the public.¹³⁴ While it is unknown how many retirements from the bench were hastened by poor judicial performance evaluations which were not publicly disseminated, such occurrences take place routinely.¹³⁵ Such instances are surely examples of substandard judges being held accountable for their performance.

The clearest example of a judicial performance evaluation program having a direct impact on a judicial election and judicial accountability occurred in 2008 in Pierce County, Washington—a state that does not use the Missouri plan to select judges but rather contested non-partisan elections.¹³⁶

In 2007, the Tacoma-Pierce County Bar Association (TPCBA) contracted to have a judicial performance evaluation conducted of the twenty-two judges sitting on the Pierce County Superior Court bench.¹³⁷ As with the JPE programs operated in states with retention elections, the Pierce County JPE was designed to provide information to the judges for self-improvement purposes and to provide information to voters to be used in considering how to cast their votes in judicial elections.¹³⁸ An additional purpose behind the bar association's conducting the JPE was to provide information to potential candidates for the superior court bench about which of the sitting judges are performing below expectations.¹³⁹

It is with the goal of fostering the potential for poorly performing judges to be held accountable that the TPCBA made two significant decisions. First, it was decided that the results of the evaluations would be released to the public four weeks before attorneys planning on entering a judicial election must notify the Secretary of State which position on a court they wish to run for.¹⁴⁰ The specific basis for this decision was to

134. See Pelander, *supra* note 112, at 685; *Editorial, Judging the Judiciary: Keep Public Informed*, DENVER POST, Nov. 9, 2007, at B-2.

135. *Judging the Judiciary*, *supra* note 134; Interview with Louise Baca-Sena, Manager, N.M. Jud. Perform. Eval., in Denver, Colo. (Aug. 6, 2008); Interview with Jane Howell, Exec. Dir., Colo. Office of Judicial Perform., in Denver, Colo. (Aug. 7, 2008).

136. See WASH. CONST. art. IV, §§ 3, 5. A caveat to this system is that if nobody files with the secretary of state to challenge a sitting judge, the judge automatically retains his or her position for another term. WASH. CONST. art. IV, § 29.

137. See Adam Lynn, *Judging the Judges: Bar Association Survey Rates Pierce County Judiciary*, THE NEWS TRIB. (Tacoma, Wash.), June 1, 2008, at A1, available at <http://www.thenewstribune.com/news/local/story/376713.html>.

138. *Id.*

139. *Bar's Judicial Ratings Will Aid the Voters*, THE NEWS TRIB. (Tacoma, Wash.), June 3, 2008, available at <http://www.thenewstribune.com/opinion/story/379136.html>.

140. E-mail from Sal Mungia, Co-chair, Tacoma-Pierce County Bar Ass'n Jud. Perform. Eval. Comm. (Apr. 24, 2008) (on file with author).

enable potential candidates to use the evaluation results to determine which of the sitting judges would be the most vulnerable in an election.

The second decision made by the TPCBA was to release to the public the written comments provided by the attorneys and jurors who evaluated the judges. This decision generated considerable controversy and consternation among the superior court's judges as well as other interested individuals and organizations in Washington. Opponents of releasing the comments, including this author, argued that individual comments about a judge can be taken out of context, have no indicia of reliability due to their anonymity, and may shift the focus from a quantitatively reliable evaluation based on established criteria toward salacious innuendo and cherry-picked, non-performance-based comments.¹⁴¹ Proponents of releasing the comments, including all members of the TPCBA judicial evaluation committee, argued that the comments provided specific and vivid examples of a judge's performance that could not be adequately conveyed by numerical ratings. In the end, the TPCBA decided to release the comments to the public on May 6, one month before the June 6 filing deadline for the election.

While the Pierce County JPE report did not provide an overall score for each judge from the four categorical indices and the responses to the twenty-five specific questions that were reported, there were two judges who were rated significantly lower than the rest of the bench.¹⁴² One of these judges was Sergio Armijo.¹⁴³ Based on attorney evaluations, Judge Armijo was the lowest rated judge in the areas of legal ability, integrity, and impartiality, and was near the bottom in the other two categories. His ratings were such that the local newspaper singled him and another judge out as being the worst judges on the bench.¹⁴⁴

In late May, Michael Hecht, an attorney from Tacoma, Washington, filed to challenge Judge Armijo for his seat on the superior court bench.¹⁴⁵ The campaigns, which ran from late May until election day on August 19, were fascinating in the dissimilarity.

As an incumbent, Judge Armijo ran a traditional, well-funded campaign. He focused on his service to the community and his experience on the bench.¹⁴⁶ He was endorsed by nearly all of the other superior

141. *Bar's Judicial Ratings Will Aid the Voters*, *supra* note 139.

142. *Id.*; Lynn, *supra* note 137. The Tacoma News Tribune developed an interactive web site to allow visitors to generate their own rankings of the judges based on their preferred criteria.

143. Lynn, *supra* note 137.

144. *Id.*

145. Judge Armijo was the only Pierce County Superior Court judge to face electoral competition in 2008. *Pierce County Superior Court Challengers Face Off*, THE NEWS TRIB. (Tacoma, Wash.), Aug. 4, 2008, available at <http://www.thenewstribune.com/news/election/story/433275.html>.

146. *Id.*

2008] *THE USE OF JUDICIAL PERFORMANCE EVALUATION* 23

court judges and the local prosecuting attorney's association.¹⁴⁷ He also raised over \$45,000 between June and August for his campaign.¹⁴⁸ These funds included nearly \$3,000 from judges sitting in Pierce County, \$400 from the Pierce County Prosecuting Attorney's Association, and several thousands of dollars from attorneys;¹⁴⁹ each of these groups is a regular source of campaign funding for incumbent judges facing an electoral challenge. Judge Armijo's campaign spent its money on direct mailings (over \$20,000), yard signs (\$8,000), and a campaign manager (over \$2,000).¹⁵⁰

His opponent, Michael Hecht, ran a much different campaign. Hecht raised only \$6,700 for his campaign, none of which came from sitting judges and less than \$1,000 of which came from attorneys.¹⁵¹ The focus of his campaign was on the need to hold Judge Armijo accountable for his poor performance. This is exemplified by the personal statement he provided to the local newspaper, the *Tacoma News Tribune*:

My opponent's dismal performance has harmed the citizens of Pierce County. Families broken and financially ruined, minorities, senior citizens and women victimized, law enforcement disrespected. It is time for a change. I will not run personal business from my court. I will not ask lawyers who practice before me for their endorsements. Fundamental fairness is a cornerstone of justice, ability to apply the law is paramount. For a fair day in court, elect Hecht Judge.¹⁵²

Similarly, when asked why he would be the best candidate for the position, he responded:

Necessity. Tacoma Pierce County Bar Association got it right, 2008 Evaluation www.tpcba.com. Judge Armijo rated last of all 22 judges in legal ability, integrity and impartiality. I pledge decisions based on applying the law; fairness and impartiality, protecting families and children, reading and being prepared, and avoiding conflicts of interest. I respect law enforcement and cultural diversity. The judicial

147. WASH. STATE PUB. DISCLOSURE COMM'N, CASH RECEIPTS MONETARY CONTRIBUTIONS (2008) [hereinafter MONETARY CONTRIBUTIONS I], available at <http://www.pdc.wa.gov/rptimg/Default.aspx?docid=1186949>; WASH. STATE PUB. DISCLOSURE COMM'N, CASH RECEIPTS MONETARY CONTRIBUTIONS (2008) [hereinafter MONETARY CONTRIBUTIONS II], available at <http://www.pdc.wa.gov/rptimg/Default.aspx?docid=1212507>.

148. WASH. STATE PUB. DISCLOSURE COMM'N, CANDIDATE SUMMARY, FULL REPORT, RECEIPTS AND EXPENDITURES, <http://www.pdc.wa.gov/rptimg/Default.aspx?docid=1221563>.

149. See MONETARY CONTRIBUTIONS I, *supra* note 147; MONETARY CONTRIBUTIONS II, *supra* note 147. This figure is from donors who had their occupation listed as attorney on campaign finance disclosure forms submitted to the Washington Public Disclosure Commission. A number of other donors were also attorneys by trade but not listed as such.

150. WASH. STATE PUB. DISCLOSURE COMM'N, EXPENDITURES CONTINUATION SHEET (2008), available at <http://www.pdc.wa.gov/rptimg/Default.aspx?docid=1192260>.

151. WASH. STATE PUB. DISCLOSURE COMM'N, CAMPAIGN SUMMARY RECEIPTS & EXPENDITURES (2008), available at <http://www.pdc.wa.gov/rptimg/Default.aspx?docid=1214451>; WASH. STATE PUB. DISCLOSURE COMM'N, CASH RECEIPTS MONETARY CONTRIBUTIONS (2008), available at <http://www.pdc.wa.gov/rptimg/Default.aspx?docid=1192049>.

152. *Pierce County Superior Court Challengers Face Off*, *supra* note 145.

process is more important than any one individual. I am committed to being a good judge.¹⁵³

This focus was made even sharper with the launching of the campaign's website, hechtforhelp.com.¹⁵⁴ Throughout the multipage website, references and links to materials pertaining to Judge Armijo's results in the Pierce County Superior Court Judicial Performance Evaluation are found. A page titled "Why I am Running" lays out repeatedly how Judge Armijo was viewed as being the worst judge in Pierce County in the performance evaluation and needed to be replaced.¹⁵⁵ Also on the site are numerous verbatim comments provided by attorneys as part of their evaluation of Judge Armijo.

In an upset, Michael Hecht defeated Judge Armijo in the August election, receiving 51% of the vote.¹⁵⁶ While exit polling was not done after the election—and given the disparity in funds and endorsements—it is safe to say that the performance evaluation was taken to heart by the electorate and served as a primary factor behind Judge Armijo's defeat.

C. *Do JPE Programs Encourage Higher Voter Participation Rates in Judicial Elections?*

As noted above, a common complaint raised by opponents of merit selection and retention elections is the lack of information provided to voters in retention elections and the low level of voter participation. Research has indicated that there is a distinct relationship between the lack of information possessed by voters regarding judicial candidates and voter turnout. Professors Lovrich and Sheldon found that citizens who voted in judicial elections possessed significantly more knowledge about judicial candidates than citizens who did not vote.¹⁵⁷ Similar results have been found in studies examining falloff and non-judicial elections.¹⁵⁸

It stands to reason that if lack of information is related to lower voter participation, then providing voters with more useful information should increase participation rates. If we accept this premise and if judicial performance evaluation programs are meeting their goal of inform-

153. *Id.*

154. Michael Hecht, *Elect Hecht: Hecht for Help Homepage*, <http://hechtforhelp.com> (last visited Oct. 21, 2008).

155. Michael Hecht, *Elect Hecht: Hecht for Help, Why I am Running*, http://hechtforhelp.com/abouthecht_why.html (last visited Oct. 21, 2008).

156. See Adam Lynn, *Incumbent 15-Year Judge Armijo Ousted in Final Election Result*, THE NEWS TRIB. (Tacoma, Wash.), Sep. 12, 2008, available at <http://www.thenewstribune.com/news/local/story/480441.html>.

157. Sheldon & Lovrich, *supra* note 88, at 238; Nicholas P. Lovrich, Jr. & Charles H. Sheldon, *Voters in Judicial Elections: An Attentive Public or an Uninformed Electorate?*, 9 JUST. SYS. J. 23, 30 (1984).

158. See Shaun Bowler, Todd Donovan & Trudi Happ, *Ballot Propositions and Information Costs: Direct Democracy and the Fatigued Voter*, 45 W. POL. Q. 559, 560-61 (1992); John E. Mueller, *Voting on the Propositions: Ballot Patterns and Historical Trends in California*, 63 AM. POL. SCI. REV. 1197, 1200 (1969); Wattenberg et al., *supra* note 86, at 236, 247.

2008] *THE USE OF JUDICIAL PERFORMANCE EVALUATION* 25

ing voters about judges standing for retention, then falloff in judicial retention elections should decrease, indicating increased voter participation.

To see if this is in fact taking place, I reviewed election turnout data in Arizona, Colorado, and New Mexico. Each of these states selects its appellate court judges via merit selection and retention elections and has a well established JPE program. The charts presented in Figures 2, 3, and 4 present a level of falloff between the number people who submitted a ballot in an election and the number of people who, on average, in a statewide judicial retention election that appeared on the ballot. Each chart has a line bisecting the horizontal axis, indicating the year in which judicial performance evaluation information about appellate court judges standing for retention was made available to the public for use in retention elections. If the proposition that increased information about judges standing for retention will lead to increased voter participation in retention elections, then the level of falloff should decrease after JPE programs are established. As is evident from the charts, falloff decreased as predicted in two states, Colorado and New Mexico, but increased in Arizona.

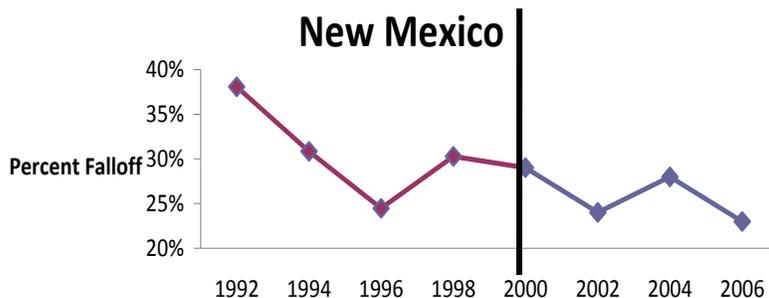


FIGURE 1

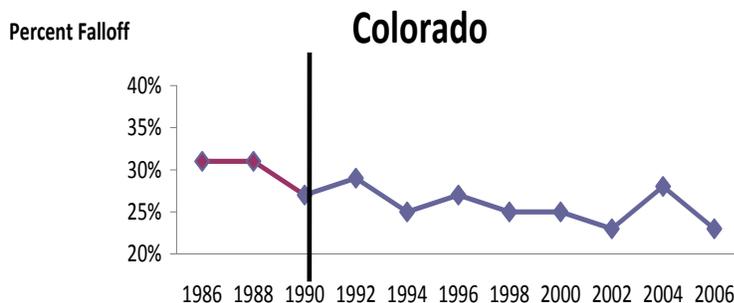


FIGURE 2

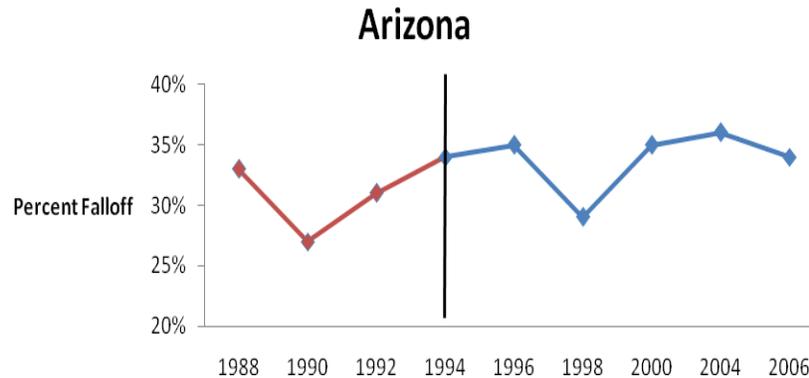


FIGURE 3

What can be gleaned from this information? First and foremost, such a cursory analysis should be considered exploratory. Not only is it impossible for one to make any causal inferences from the results, but there are also a number of items that need to be controlled for statistically before one can make any concrete findings about the effect of JPE programs on voter participation.¹⁵⁹ While preliminary, the information does suggest that JPE programs, in and of themselves, are not likely to be sufficient to eliminate voter falloff.

D. How do Judges Perceive the Impact of JPE on Judicial Accountability and Judicial Independence?

A critical component of the ability of JPE programs to promote judicial accountability without negatively impacting judicial independence is how the judges under evaluation perceive the program. Given the sensitive nature of judicial performance evaluation programs in the eyes of judges subject to evaluation, it is important to obtain input on the design, implementation, and assessment of JPE programs from the judges being evaluated.¹⁶⁰ Input from the judiciary is likely to make for smoother implementation and more reliable results than if judicial feedback is not considered.¹⁶¹ Moreover, by considering the thoughts and concerns of the judges, judges will attain “ownership” of the process which would

159. State and election-specific items such as methods of distributing evaluation recommendations, population transiency, ballot design, length of ballot, and other items must be taken into account before definitive findings and attributions can be made.

160. Daina Farthing-Capowich, *Designing Programs to Evaluate Judicial Performance*, 9 ST. CT. J. 22, 23-24 (Summer 1985).

161. *Id.*

2008] *THE USE OF JUDICIAL PERFORMANCE EVALUATION* 27

help quell potential concerns raised by judges who are rated more critically than they feel is warranted.¹⁶²

There are several methods that can be used to gather this information from judges. Individual judges can be interviewed in private regarding their perceptions of a JPE program. Interviews have the ability to obtain detailed information and thoughts from judges that can provide much insight. In practicality, however, to interview a representative sample of judges in a state would be a daunting task. Such interviews would likely be very time consuming, expensive, and logistically cumbersome.¹⁶³

A second option would be to conduct a series of focus groups with members of the bench. This would entail meeting with groups of judges and having structured discussions of their perceptions about JPE programs. While this is less time consuming and expensive than individual interviews, it does not provide the confidential atmosphere present with a one-on-one interview setting. It is not unrealistic to expect judges to be less candid in discussing their thoughts about evaluations of the judiciary in front of their fellow judges than they would be in a private setting. Accordingly, the information obtained in focus groups may be incomplete and skewed based on the dynamics of the group.

A third method of gaining insight into what judges think about JPE programs is by conducting a survey. Surveys have the advantages of providing anonymity, being able to obtain quantifiable responses that can be examined for an entire judiciary using standardized scales, and not being cost prohibitive. Additionally, in asking open-ended questions, surveys can contain significant levels of detail and opinions of individual judges.

The limited number of surveys of judges who have been evaluated by JPE programs have revealed interesting attitudes about JPE's effect on judicial independence and accountability. As part of its multistate examination of JPE programs, the American Judicature Society (AJS) surveyed the judges from Alaska, Arizona, Colorado, and Utah who were evaluated by their state's JPE program in 1996 about their perceptions of the programs.¹⁶⁴ AJS found that a significant majority of those judges surveyed believed the JPE programs in their states helped make them "appropriately accountable for [their] job performance."¹⁶⁵ According to

162. *Id.*; See Gary E. Roberts, *Employee Performance Appraisal System Participation: A Technique that Works*, 31 PUB. PERSONNEL MGMT. 333, 334 (2002) (noting that employees are more likely to accept negative feedback if they believe the evaluation process is fair).

163. FRANK H. HAGAN, *RESEARCH METHODS IN CRIMINAL JUSTICE AND CRIMINOLOGY* 186 (7th ed. 2006).

164. Esterling, *supra* note 35, at 211-14.

165. *Id.* at 211.

one judge, “the public is better informed when making voting decisions and can weed out incompetent judges.”¹⁶⁶

Recently, members of the Colorado judiciary participated in a survey which sheds a great deal of light on the Colorado JPE process.¹⁶⁷ In 2008, the author of this article, in conjunction with Jordan Singer of the Institute for the Advancement of the American Legal System, conducted a study to gain insight into what members of the Colorado judiciary thought about Colorado’s system of judicial performance evaluation. As part of the study, surveys inquiring about judicial perceptions of the JPE program were sent to Colorado’s 270 district and county court judges as well as the state’s twenty-six appellate court judges.¹⁶⁸ In all, 172 (64%) trial court judges and seventeen (65%) appellate court judges returned completed surveys.

While the survey did not ask the judges specifically about whether the JPE process affects judicial accountability, a number of judges raised the issue *sua sponte* in their written comments submitted as part of the survey. A number of judges noted that JPE makes them more accountable to the public. For example, one county court judge noted, “My belief is that the process requires me to be more accountable than I would be without it. The process helps prevent ‘black robe disease.’”¹⁶⁹ Similarly, a district court judge wrote, “I think the existence of the process is useful. Knowing that you will be evaluated is a good hedge against judicial arrogance. I also think the process weeds out some appointees who don’t turn out to be well suited for the job.”¹⁷⁰

From these and similar comments, it appears that Colorado judges perceive that the JPE process increases their accountability to the public. Furthermore, this increased accountability appears to have a positive impact on judicial behavior. By helping judges identify occasions where they may be suffering from “judicial arrogance” and “black robe disease,” the accountability associated with the performance evaluation process appears to operate beyond electoral defeat for the betterment of the judicial system.

In conjunction with considering JPE’s impact on judicial accountability, it is worthwhile to examine its effect on judicial independence. Recall, the establishment of JPE programs was based largely on the premise of preserving judicial independence by increasing the ability of

166. *Id.*

167. David Brody & Jordan Singer, 2008 Survey of the Colorado Judiciary About Judicial Performance Evaluation, Mar. 26, 2008 [hereinafter Colorado Judicial Survey] (on file with author).

168. *Id.* (response data on file with author). See also INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., THE BENCH SPEAKS ON JUDICIAL PERFORM. EVAL.: A SURVEY OF COLORADO JUDGES (2008), available at <http://www.du.edu/legalinstitute/publications2008> (for a presentation of an executive summary of the findings).

169. Colorado Judicial Survey, *supra* note 167 (comment of Trial Judge 15).

170. *Id.* (comment of Trial Judge 27).

2008] *THE USE OF JUDICIAL PERFORMANCE EVALUATION* 29

voters to use retention elections to hold judges accountable based on appropriate, non-decision-based criteria.¹⁷¹ An evaluation program that fails to adequately preserve judicial independence is likely to lose the support of a large segment of individuals and organizations that worked for its establishment.

The 2008 survey of the Colorado judiciary asked the judges whether the JPE program increases judicial independence, decreases it, or has no effect on it. As can be seen in Figure 5, it is fair to say that the judges were evenly split on what the impact of JPE is on judicial independence.

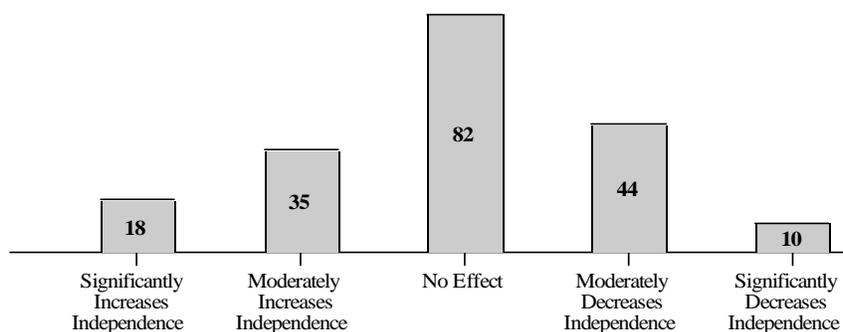


FIGURE 5
COLORADO JUDGES' OPINIONS ON EFFECT OF JUDICIAL PERFORMANCE
EVALUATIONS ON JUDICIAL INDEPENDENCE

When one considers the written comments that were provided by the judges relating to judicial independence, several items become evident. First, judges who believe JPE increases independence viewed it as a necessary component of Colorado using the Missouri Plan to select judges. As summed up by a district court judge, "Without JPE and its . . . reports to the voters, the retention system would collapse. Therefore, JPE has a significant impact on judicial independence."¹⁷² Similarly, another district court judge stated, "I believe that JPE increases judicial independence because I think it is the price we must pay to continue to have a merit selection system for judges."¹⁷³

As for those who believed JPE decreased judicial independence, no judge provided information about how it did so. Rather, much of the concern about its effects on independence centered on procedural issues associated with the survey process. The importance of considering and

171. See *supra* Section IV.

172. Colorado Judicial Survey, *supra* note 167 (comment of Trial Judge 4).

173. *Id.* (comment of Trial Judge 52).

evaluating the procedures used in conducting a JPE program and the effect the procedures have on perceptions of the effectiveness and trustworthiness of a program are discussed below.

E. Are Proper Processes Being Used to Evaluate Judicial Performance?

“The effectiveness of any judicial performance evaluation project will depend, in large measure, upon the reliability of the information it generates.”¹⁷⁴ This statement, included in the American Bar Association’s first set of judicial performance evaluation guidelines over twenty years ago, is as true today as it was when written. For judicial performance evaluations to effectively hold judges accountable while not inhibiting judicial independence, the results of the evaluations and the methods used to obtain them must be unquestionably trustworthy.¹⁷⁵

The methods used and the recommendations and reports generated must be viewed as being trustworthy by the judges being evaluated, the voters who are being asked to use them, the state commissions and overseers who must stand behind them, and state policy makers who make funding decisions.¹⁷⁶ A lack of trust in the methods selected to obtain the information used in the evaluation process by any of these groups is likely to have immediate and long-term consequences for the program. As Seth Andersen, Executive Director of the American Judicature Society put it:

While the current trend appears to favor adoption of official retention evaluation programs in more states, it is important to note that only six of the nineteen states that hold retention elections at some or all levels of court have adopted such programs. Concerns about the fairness of survey methodologies and evaluation commission procedures, as well as a general reticence among many judges to subject themselves to an evaluation process that may be seen as a threat to decisional independence, have helped to stall the expansion of retention evaluation programs.¹⁷⁷

Despite the logic behind Andersen’s observation, there have not been any external evaluations of the survey methodologies used by JPE programs since the 1990s. Given the importance the trustworthiness and validity of a program’s methods and results have on the effectiveness of JPE programs in promoting judicial accountability, it is important that states routinely assess the survey methodologies, commission proce-

174. SPECIAL COMM’N ON EVALUATION OF JUDICIAL PERFORMANCE, AMERICAN BAR ASSOCIATION GUIDELINES FOR THE EVALUATION OF JUDICIAL PERFORMANCE 26 (1985).

175. *Id.* See also ABA, *supra* note 40, at 6.

176. See N.C. BAR ASS’N JUDICIAL PERFORM. EVAL. STUDY GROUP, REPORT AND RECOMMENDATIONS OF THE NCBA JPE STUDY GROUP ON JUDICIAL PERFORM. EVAL TO THE N. C. BAR ASS’N BOARD OF GOVERNORS 27 (2006), available at <http://www.ncbar.org/download/ncba/jpeDraftReport.pdf>.

177. Andersen, *supra* note 35, at 1375-76 (footnote omitted).

dures, and distribution of results being used, before negative perceptions of their fairness undermine public and political acceptance. The import of such assessments is evident from the activities and events surrounding Colorado's judicial selection and performance evaluation process over the past several years. The following section presents a truncated case study of JPE in Colorado.

V. JUDICIAL PERFORMANCE EVALUATION IN COLORADO

Colorado has one of the longest running and most well respected judicial performance evaluation programs in the nation.¹⁷⁸ Established by the Colorado legislature in 1988, the Commissions on Judicial Performance conduct evaluations of each appellate and trial court judge prior to the end of the judge's term.¹⁷⁹ The statutes and rules governing the processes and criteria used in the evaluation of judges lay out detailed requirements regarding what the commissions are to consider in conducting their evaluations and how the information should be reported to the public.¹⁸⁰

As discussed previously, in 2008 I collaborated with the Institute for the Advancement of the American Legal System at the University of Denver to conduct a survey of the members of the Colorado judiciary about their thoughts regarding the JPE program. The judicial survey was the first part of a multistep plan to evaluate the processes used and impact resulting from the JPE program. To help lay the groundwork for future areas of inquiry, the judicial survey was designed to be exploratory and to garner as much information about the process as possible from the people directly affected by the JPE program.

According to Daina Farthing-Capowich, a forerunner in the design of JPE programs, in developing and assessing the effectiveness and validity of a performance evaluation program it is important that judges be given the chance to "vent frustrations, comment, offer suggestions, and review the work product as it takes shape."¹⁸¹ Input from the subjects of a performance evaluation gives insight into matters that are worthy of assessment and contemplation and provides the impetus in the discovery of systemic shortcomings that can be easily remedied.¹⁸² The 2008 Colorado judicial survey, discussed above, is a classic example of this phenomenon.

178. Jean E. Dubofsky, *Judicial Performance Review: A Balance Between Judicial Independence and Public Accountability*, 34 *FORDHAM URB. L.J.* 315, 315 (2007).

179. Commissions on Judicial Performance, *COLO. REV. STAT. ANN.* §§ 13-5.5-101 to -106 (West 2008). Colorado has 22 local commissions which evaluate trial court judges and a state commission that evaluates supreme court and court of appeals judges.

180. *COLO. REV. STAT. ANN.* § 13-5.5-105.5 (West 2008).

181. Farthing-Capowich, *supra* note 160, at 24.

182. *Id.*

The judges' responses to the survey showed that nearly all of them were very supportive of the continued use of judicial performance evaluations in Colorado. Over 70% of judges who had been through the process at least once reported that it provided them with information that allowed them to improve their job performance, and nearly 90% believed the Commissions' recommendation process for use in retention elections was fair.¹⁸³ They did, however, have a number of strong opinions about the validity of the survey methodology and how the evaluation process could be improved. While much of this information is critical of the process, it also provides important information for the Commission to consider. Given the Colorado Commission's internal assessment of the evaluation process following each round of retention elections, and its support and assistance in the implementation of the survey, the feedback provided by the bench and discussed below will be used to make a model JPE program even stronger.¹⁸⁴

A. Perceived Problems with the Survey Methodology

To help us learn about how judges feel about JPE, the 2008 survey asked the judges whether certain aspects of the evaluation process were problematic to them. As can be seen from the figures in Table 1, most of the judges considered specific procedural items to be of major concern. Specifically, a clear majority of the judges felt that the number of respondents, the manner in which respondents are selected, and the groups of individuals identified to participate were a problem. Perhaps most importantly, only 15 judges (12.3%) responded that the validity and accuracy of the survey responses were not a problem with the Colorado JPE system.

TABLE 1: PERCEIVED PROBLEMS WITH JPE SURVEY PROCESS

	Not a problem	Minor problem	Major problem
How job performance criteria are measured.	32.1%	49.5%	18.3%
The targeted survey respondent groups.	31.0%	36.2%	32.8%
Number of survey respondents.	13.5%	33.3%	53.2%
The methods by which respondents are selected.	28.9%	40.2%	30.9%
Validity & accuracy of survey responses.	13.3%	46.0%	40.7%

As striking as these figures are, the magnitude of the dissatisfaction felt by the judges regarding the evaluation methodology is better illustrated from the following comments that were provided by a number of judges.

183. Colorado Judicial Survey, *supra* note 167.

184. Interview with Jane Howell, *supra* note 135.

- Survey is unscientific and a joke.¹⁸⁵
- Asking non-lawyers to address questions of legal knowledge is problematic—there is no legitimate way to know whether the respondents have the requisite knowledge or adequate information to support their ratings.¹⁸⁶
- The surveys are not statistically valid and are totally unreliable.¹⁸⁷
- The evaluation information is based on a survey process that has been inherently flawed, thereby limiting the value of any information we receive.¹⁸⁸
- I have now been on both sides of this evaluation process. Many years ago I was a member of our local judicial performance commission. Back then the methodology of the process of obtaining information from attorneys, witnesses, jurors was more than flawed. The survey methodology astoundingly remains horrible and unreliable. I am shocked that [the contractor] has not improved its process. . . . When a judge's career may hinge on these surveys and the opinion of the commission it is not acceptable to me that the methodology is flawed. All judges deserve better.¹⁸⁹
- The surveys are a major concern for many judges, including me. For a survey to be valid it must be provided to a wide group of people in various categories, including prosecutors, defense attorneys, court staff, probation officers, police officers, and pro se parties. Many of the results I have seen have a significant number or group that is underrepresented.¹⁹⁰

To better understand the nature of these concerns, and in an effort to see if there were any areas where the JPE process could be strengthened, I examined the 2008 judicial performance evaluation reports prepared for the Colorado trial court judges standing for retention elections. Three major items involving the evaluation of county and district court judges stood out that should be considered by the Colorado commission.

185. Colorado Judicial Survey, *supra* note 167 (comment of Trial Judge 14).

186. *Id.* (comment of Trial Judge 1).

187. *Id.* (comment of Trial Judge 18).

188. *Id.* (comment of Appellate Judge 1).

189. *Id.* (comment of Trial Judge 134).

190. *Id.* (comment of Trial Judge 82).

B. Survey Sample Issues

A key component of JPE programs is their obtaining evaluation information from individuals who actually appeared before a judge. In Colorado, this information is generated through government records and provided to the consulting firm who administers the JPE surveys for the Commissions on Judicial Performance. These individuals include attorneys, witnesses, jurors, litigants, and victims.¹⁹¹

A concern raised repeatedly by the Colorado judges was the small number of attorneys who completed surveys and what the makeup of the attorney sample looked like. In large jurisdictions, an attorney may have appeared before multiple judges during an evaluation period. Out of concern for potential “survey fatigue,”¹⁹² Colorado has adopted a unique policy whereby attorneys are only asked to evaluate up to two judges before whom they have appeared rather than ask all attorneys who appeared before a judge to complete a survey.¹⁹³ If an attorney appeared before more than two judges being evaluated by the Commission during an evaluation period, the consultant who administers the evaluation survey process draws a small sample of the attorneys who appeared before a judge, and selects those attorneys to receive evaluations for the judge.¹⁹⁴ While the sampling is based on the number of times an attorney appeared before a judge, it ignores the makeup of the sample and how it may affect the outcome of the evaluation or perception of fairness.¹⁹⁵

Ideally, a sample would include a weighted cross-section of respondents stratified by key characteristics.¹⁹⁶ In the case of a JPE program, any sampling should consider an attorney’s general area of practice (civil or criminal), and whether the attorney is a district attorney/prosecutor, or

191. COLO. REV. STAT. ANN. § 13-5.5-101-13.5.5-101.5 (West 2008); COLO. RULES GOVERNING THE COMM’NS ON JUDICIAL PERFORM., R. 10(a), 11(a).

192. While survey fatigue or response burden is a concern in asking individuals to complete multiple surveys or evaluations, other JPE programs have attorneys complete many more than two evaluations during an evaluation period without response shortfalls due to survey fatigue. Moreover, as attorneys in the Denver area are likely to appear before fewer than fifteen judges under evaluation in a given year, as opposed to fifty-two in King County, Washington and over seventy in Maricopa County, Arizona, survey fatigue should not be an issue. See JUDICIAL EVAL. COMM., KING COUNTY BAR ASS’N, 2007 JUDICIAL EVAL. SURVEY 1 (2007), available at <http://www.kcbar.org/judicial/pdf/2007judicial.pdf>; Arizona Comm’n on Judicial Perform. Rev., Judicial Perform. Reports, <http://azjudges.info/reports/lastname.cfm> (last visited Oct. 21, 2008).

193. The methodology used in the Colorado Judicial Performance Evaluation program is discussed in the individual judge reports. For ease of reference, I will provide the web page and report page number for one report when discussing system wide methods. For explanation of sampling methods and survey fatigue, see COMM’N ON JUDICIAL PERFORM., MARTIN F. EGELHOFF 2008 JUDICIAL PERFORM. SURVEY, 91 (2008), available at http://www.cojudicialperformance.com/images/retentionpdfs/2008_Dst%2002%20Martin%20F.%20Egelhoff.pdf.

194. *Id.* This method of selecting who an attorney may evaluate rather than letting attorneys select what judges they feel capable of evaluating has the potential to severely curtail response rates and levels.

195. *See id.*

196. Edward L. Korn & Barry I. Graubard, *Examples of Differing Weighted and Unweighted Estimates from a Sample Survey*, 49 AM. STATISTICIAN 291, 291 (1995).

2008] *THE USE OF JUDICIAL PERFORMANCE EVALUATION* 35

a criminal defense attorney. In addition to concerns about the small number of attorneys involved in the JPE process, it is this sampling process that is the cause of the judges' complaints.

To illustrate the nature of these problems, Table 2 contains the attorney samples used for the 2008 performance evaluations for District Court judges sitting in Denver.¹⁹⁷ First, consider the overall number of attorneys included in the sample and completing evaluations. Despite the fact that each of the district court judges sitting likely has hundreds of attorneys appear before him or her annually, on average only seventy-four attorneys received requests to complete surveys per judge. Furthermore, an average of only thirty-four attorneys completed evaluations per judge. Having only thirty-four attorneys evaluate a judge is not an *unreasonably* low figure, but the goal of any JPE program should be to have as great a number of respondents as possible.¹⁹⁸ To achieve this goal, one should not deliberately use small sample sizes when they are not necessary. This is especially true for urban areas with where large numbers of attorneys are likely to have appeared before each judge.¹⁹⁹

TABLE 2 2008 ATTORNEY EVALUATION PARTICIPANTS, COLORADO DISTRICT COURT JUDGES, DISTRICT 2

Judge	District Attorney		Criminal Defense		Civil Attorneys		Attorneys (Unknown Role)		TOTAL ATTORNEYS	
	Sent	Complete	Sent	Complete	Sent	Complete	Sent	Complete	Sent	Complete
Egelhoff	28	14	25	14	11	4	35	15	99	47
Hyatt	27	18	3	0	34	17	9	4	73	39
Madden	3	2	45	20	9	0	0	0	57	22
Mansfield	1	0	51	15	4	2	13	4	69	21
McGahey	13	5	8	3	43	25	9	4	73	37
Naves	13	9	2	1	6	3	16	6	37	19

197. Information used in this analysis was obtained from the 2008 individual judge reports prepared by the Colorado Commission on Judicial Performance: COMM'N ON JUDICIAL PERFORM., ANNE MANSFIELD 2008 JUDICIAL PERFORM. SURVEY, 93 (2008), *available at* http://www.cojudicialperformance.com/images/retentionpdfs/2008_Dst_02_Anne_Mansfield.pdf; COMM'N ON JUDICIAL PERFORM., JOHN MADDEN 2008 JUDICIAL PERFORM. SURVEY, 93 (2008), *available at* http://www.cojudicialperformance.com/images/retentionpdfs/2008_Dst_02_John_Madden.pdf; COMM'N ON JUDICIAL PERFORM., LARRY J. NAVES 2008 JUDICIAL PERFORM. SURVEY, 93 (2008), *available at* http://www.cojudicialperformance.com/images/retentionpdfs/2008_Dst_02_Larry_J_Naves.pdf; COMM'N ON JUDICIAL PERFORM., MARTIN F. EGELHOFF 2008 JUDICIAL PERFORM. SURVEY, *supra* note 193, at 93; COMM'N ON JUDICIAL PERFORM., ROBERT S. HYATT 2008 JUDICIAL PERFORM. SURVEY, 93 (2008), *available at* http://www.cojudicialperformance.com/images/retentionpdfs/2008_Dst_02_Robert_S_Hyatt.pdf; COMM'N ON JUDICIAL PERFORM., ROBERT L. MCGAHEY, JR. 2008 JUDICIAL PERFORM. SURVEY, 93 (2008), *available at* http://www.cojudicialperformance.com/images/retentionpdfs/2008_Dst_02_Robert_L_McGahey_Jr..pdf; COMM'N ON JUDICIAL PERFORM., SHEILA ANN RAPPAPORT 2008 JUDICIAL PERFORM. SURVEY, 93 (2008), *available at* http://www.cojudicialperformance.com/images/retentionpdfs/2008_Dst_02_Sheila_Ann_Rappaport.pdf.

198. See JUDICIAL EVAL, COMM., *supra* note 192, at 3.

199. In rural counties, low numbers of attorneys completing surveys may be unavoidable to the small population of attorneys practicing locally. This is generally not the case in urban locales such as Denver.

Rappaport	20	13	0	0	32	20	9	6	61	39
-----------	----	----	---	---	----	----	---	---	----	----

TABLE 3 2008 NON-ATTORNEY EVALUATION PARTICIPANTS,
COLORADO DISTRICT COURT JUDGES, DISTRICT 2

Judge	Law Enforcement		Criminal Defendant		Other		Civil Litigant		Jurors	
	Sent	Complete	Sent	Complete	Sent	Complete	Sent	Complete	Sent	Complete
Egelhoff	161	18	116	5	141	20	38	1	284	107
Hyatt	4	0	99	3	21	0	10	0	137	84
Madden	1	0	195	9	0	0	1	0	170	71
Mansfield	0	0	239	13	119	15	85	25	143	53
McGahey	113	7	77	5	147	9	20	1	247	126
Naves	1	1	14	0	37	1	0	0	184	102
Rappaport	5	0	45	4	8	0	1	0	272	154

Of greater concern with the samples used for the Denver judges is the professional makeup of the attorneys. Ideally there should be a reasonable balance between the number of prosecutors and defense attorneys who evaluate a judge. While this is the case for two of the Denver judges, Egelhoff and McGahey, the same cannot be said for the five other judges. For three of the judges, surveys were sent to more than a dozen district attorneys while three or fewer defense attorneys were included. For the other two judges the opposite is true. For Judges Madden and Mansfield, surveys were sent to forty-five and fifty-one criminal defense attorneys, but to only three and one district attorneys, respectively. Given that a judge who hears criminal matters is likely to have both prosecutors and defense attorneys appear in his or her court, this disparity is puzzling and of understandable concern to the bench.

The same is true for the distribution of surveys to witnesses and litigants. As shown in Table 3, there was great disparity in the numbers of surveys sent to law enforcement officers for each judge. For two judges over 100 surveys were sent to law enforcement personnel, while for the other five judges five or fewer were sent to police officers per judge. At the same time, judges who had the opportunity to be evaluated by five or fewer law enforcement officers had surveys sent to dozens of criminal defendants each.

The fact that it is done intentionally is problematic; however, improved perceptions of the reliability and trustworthiness of the process can be obtained when these items and the judges' concerns about them are addressed in future evaluations.

C. Use of a Single Survey for Non-Attorney Evaluators

The second area of concern raised by the judges involves the non-attorney survey instrument that is used in the JPE process. The reason behind having lay persons evaluate judges as part of a judicial performance evaluation is that they bring different perspectives into the assessment process.²⁰⁰ To make use of the different perspectives it is important that a survey questionnaire be specifically tailored to the nature of the interaction each group of respondents had with a judge.²⁰¹ While all lay persons have some common interactions with a judge, the nature of sitting on a jury for several days or weeks makes the experience and perspective it provides inherently different than that of a witness who testified for one hour. For this reason, all state JPE programs have evaluation questionnaires tailored specifically for jurors, apart from those used with witnesses and litigants. This gives respondents an opportunity to provide in-depth evaluations about the judge involving certain events and observations that are not relevant or applicable to other lay persons.

Under the Colorado JPE process, all non-attorneys complete the same evaluation questionnaire. The practice of having jurors, litigants, criminal defendants, social workers, law enforcement officers, and victims complete the same survey limits the value of the information they can provide. Compounding the matter, when evaluation reports are prepared for each judge, the responses of all non-attorney respondents are pooled together and reported as one measure.²⁰² Given the fact that in the aggregate jurors may perceive a judge's behavior differently than law enforcement officers who evaluate a judge, a judge's rating could potentially be increased or decreased depending on the number of each type of lay person that complete an evaluation.²⁰³

200. ABA, *supra* note 40, at 14; Esterling, *supra* note 35, at 210.

201. See SHARED EXPECTATIONS, *supra* note 34, at 65.

202. COMM'N ON JUDICIAL PERFORM., *supra* note 193, at 92.

203. This problem is exasperated by the decision to rank judges based largely on these evaluation scores. See Matt Masich, *Judge Performance Evals To Be Released At 1 P.M.*, L. WK. COLO. ONLINE, Aug. 5, 2008, available at <http://www.lawweekonline.com/default.asp?sdetail=6673>. For the rankings, see COMM'N ON JUDICIAL PERFORM., COMBINED OVERALL AVERAGE & RANK 1-2 (2008), available at <http://www.cojudicialperformance.com/CO%20Web%20Ranking%20%2008-04-08a.pdf>.

D. *Non-Attorneys Rating Legal Ability*

While one of the most important aspects of judicial performance evaluations is their ability to provide information about a judge's performance as seen from lawyers and non-lawyers alike, it is essential that both groups evaluate judges only on areas in which they are competent.²⁰⁴ Of particular import is having only individuals with legal training and experience evaluate a judge's legal ability.²⁰⁵ Under the Colorado JPE system, not only do all non-attorneys complete the same survey, but they are also asked to evaluate the judge on matters involving his or her "application of the law." Specifically, jurors, litigants, witnesses, and other non-attorneys are asked to rate trial court judges on three criteria under the heading "Application of Law":²⁰⁶

1. Giving reasons for rulings.
2. Willing to make decision without regard to outside pressure.
3. Being able to identify and analyze facts.

While such questions are appropriate for attorneys to answer, expecting non-attorneys to be able to intelligently assess whether a judge applied the law appropriately is unacceptable, and justifiably troubling to the bench.²⁰⁷ The problems with the information obtained from non-attorneys are further amplified by having non-attorneys rate the judge's *sentencing practices*. The Colorado survey explicitly asks non-attorneys to assess whether the judge's sentencing practices are generally too harsh or too lenient. This is unacceptable for several reasons. First, rather than focus on a judge's behavior, this asks for assessments on decision-based criteria. More importantly, it asks people to rate a judge on information for which they have no basis of knowledge.²⁰⁸ Jurors are not present at a criminal sentencing, and therefore have no basis to provide a rating. The same is true for most witnesses and other non-attorneys who appear before a judge. While criminal defendants may be present at sentencing, many will not have a basis to evaluate the judge's relative harshness.

204. Daina Farthing-Capowich & Judith White McBride, *Obtaining Reliable Information: A Guide to Questionnaire Development for Judicial Performance Evaluation Programs*, 11 ST. CT. J. 5, 7 (Winter 1987).

205. *Id.*

206. COMM'N ON JUDICIAL PERFORM., *supra* note 193, at 98.

207. Beyond the appropriateness of having non-lawyers rate legal ability, the fact that jurors, witnesses, litigants, and the like are frequently not privy to bench and in chambers conferences and rulings, not a complete view of all of the facts at issue in a matter before the court. As such, any evaluation on such matters is pure conjecture.

208. A rudimentary premise in the design of any performance evaluation program is that evaluations by subjects who have not had the opportunity to observe the behavior they are evaluating is a major source for unreliability in performance measurement, and must be avoided. See Bryant F. Nagle, *Criterion Development*, 6 PERSONNEL PSYCHOL. 271, 277 (1953); Hannah R. Rothstein, *Interrater Reliability of Job Performance Ratings: Growth to Asymptote Level With Increasing Opportunity to Observe*, 75 J. APPLIED PSYCHOL. 322, 322 (1990).

2008] *THE USE OF JUDICIAL PERFORMANCE EVALUATION* 39

Having non-attorneys address sentencing practices can do nothing but negatively affect the creditability of the JPE process.

These concerns expressed by the Colorado judiciary have the potential to have a negative impact on the future of the state's judicial performance program. Recall, in 2006, Colorado voters were asked to consider a ballot initiative that was anti-judiciary. The initiative was born out of a belief in some quarters that the Colorado judiciary was unaccountable to the public with or without a JPE program.²⁰⁹

Critics of the Colorado Commissions on Judicial Performance believe that it fails to effectively hold judges accountable.²¹⁰ They cite the fact that, since 1988, fourteen trial court judges and no appellate court judges have received "do not retain" recommendations.²¹¹ During this time period six trial court judges have not been retained by voters and no appellate court judges have been voted off the bench.²¹²

Due to what was perceived as a lack of judicial accountability, particularly among appellate court judges, in 2006 a political committee titled Limit the Judges was established. The goal of this group was to remove sitting appellate court judges by establishing retroactive term limits. The group's efforts led to Amendment 40 appearing on the November 2006 general election ballot. Amendment 40 provided for retroactive term limits of ten years for Colorado Supreme Court and Court of Appeals judges.²¹³ Proponents of the amendment argued that the current system of retaining judges, including the Colorado Commissions on Judicial Performance, serves as a rubber stamp for judges seeking another term in office.²¹⁴

Opponents of the amendment argued that it is foolhardy to deprive the state of its most experienced jurists arbitrarily. They also argued that judicial performance evaluations do work to hold judges accountable, but acknowledged that the system needs be improved.²¹⁵ After a bitter campaign in which over one million dollars was spent, the Amendment was defeated by a 57% to 43% tally.²¹⁶

209. See John Andrews, *Judges Coddled by Sweetheart Process*, DENVER POST, Feb. 18, 2007, at E3.

210. See *id.*

211. An additional ten judges have received recommendations of "no opinion." For more information, see COMM'N ON JUDICIAL PERFORM., JUDICIAL PERFORM. FACT SHEET 3 (2008), available at <http://www.cojudicialperformance.com/2008%20fact%20sheet1.pdf>.

212. *Id.*

213. COLO. SEC'Y OF STATE, PROPOSED CONST. AMEND. FOR 2006 BALLOT (2006), <http://www.elections.colorado.gov/WWW/default/Initiatives/Title%20Board%20Filings/Final%20T%20ext%2090.pdf>.

214. See John Andrews, *The Case for Judicial Term Limits*, COLO. LAW., Feb. 2008, at 44, 46.

215. See Bruce Finley, *Performance Reviews Proposed for Judges*, DENVER POST, Oct. 3, 2006, at B4.

216. Colorado Cumulative Report, <http://www.sos.state.co.us/pubs/electionresults2006G> (last updated Dec. 13, 2006).

Limit the Judges and their allies did not take this loss lying down. After their defeat, the forces vowed to ask the voters in 2008 to vote to adopt a refined version of Amendment 40.²¹⁷ In response to this threat, opponents of term limits and supporters of a strong, independent judiciary went on the offensive to assure the public that the state's judges are and will be increasingly accountable. As part of this effort, not only was the use of judicial performance evaluation in Colorado emphasized, but acknowledgements of the need to strengthen and improve the evaluation system were made.²¹⁸

In 2008, the Colorado legislature and State Commission on Judicial Performance did enact several reforms aimed at increasing the openness of the evaluation process and the improved ability of voters to hold poorly performing judges accountable. Several public and community hearings were held by the legislature in which representatives of the state bar, civic organizations, and other interested parties spoke in support of the renewal of the JPE program. The Commissions on Judicial Performance were reauthorized by the legislature. Importantly, representatives from the judiciary also testified in support of retaining judicial performance evaluation. With a comfortable margin, the Commissions on Judicial Performance were reauthorized in the spring of 2008.²¹⁹ Reforms enacted as part of the reauthorization bill included the implementation of a notice and comment period for enactment of rules governing the evaluation process and additional evaluation criteria provided for. None of these reforms generated much controversy.

In a more drastic move designed to increase the accountability of judges, the Colorado Commissions on Judicial Performance chose to not only release retention recommendations for judges standing for retention election in 2008, but to also rank the trial court judges based on an average of the ratings attained in attorney and non-attorney evaluation questionnaires.²²⁰ The premise behind this move was that by presenting evaluation results in this manner, it would be possible that judges who scored relatively poorly in the evaluation surveys, but were judged worthy of retention by the commissions, could still be held accountable by the voters.

This effort may well increase judicial accountability. It may also, however, have a negative effect on the judiciary's continued support of the JPE process. An unscientific survey found that judges were over-

217. Limit the Judges: Campaign for Judicial Term Limits, Judicial Term Limits Target 2008 Ballot, <http://www.limitthejudges.com> (last visited Oct. 20, 2008).

218. Rebecca Love Kourlis, *Facing a Need for Reform: More Openness Would Aid Courts*, DENVER POST, Nov. 26, 2006, at E1.

219. The Colorado Index: Illuminating Colorado Issues, http://thecoloradoindex.typepad.com/the_colorado_index/2008/03/senate-bill-054.html; SB 08-054, 66th Gen. Assem., 2nd Reg. Sess. (Colo. 2008).

220. COMM'N ON JUDICIAL PERFORM., *supra* note 203.

whelmingly opposed to the new ranking system.²²¹ Given the move to increase accountability, and the judiciary's frustration with the methods used in conducting the evaluations, judicial support for the performance evaluations may begin to erode if their concerns are not taken into account.²²²

CONCLUSION

The selection of state judges in the United States is riddled with contradictory values and preferences. The public wants a fair and impartial judiciary that decides cases according to the law independent of political or professional consequences.²²³ At the same time Americans want the ability to hold judges accountable should they abuse their power or betray the public trust.²²⁴ While they prefer being able to hold judges electorally accountable, a large segment of voters, when given the opportunity to vote in judicial elections opt not to participate due to not having the information necessary to make an informed decision in a judicial election.²²⁵ When provided information about judges via expensive, vigorous campaigns, voters express concerns about special interest groups "buying judges" and the negative effect campaigning and having to raise money has on judicial independence and impartiality.²²⁶ While these contradictions provide great challenges for states to overcome, they are not insurmountable.

Any efforts to address these issues must center on trust. While the public wants an independent judiciary, its general lack of trust in governmental institutions and democratic ideals has led to the wide use of judicial elections and expensive campaigns and an informational vacuum. Judicial performance evaluations have the ability to provide information that can be used to hold judges accountable and limit the need for special interest groups and issue based campaigns to fill the information void. For JPE programs to achieve this aim it is necessary that the methods used and the product produced are deemed trustworthy, reliable, and valid by the public and the judiciary. Put another way, JPE programs must be accountable before voters will fully use them to hold judges accountable.

221. Masich, *supra* note 203.

222. As the Colorado Commission is committed to providing a system the judges can trust, in all likelihood, the Colorado State Commission will consider the thoughts of the judiciary in planning for the future. Interview with Jane Howell, *supra* note 135.

223. Bert Brandenburg, *Seizing the Accountability Moment: Enlisting Americans in the Fight to Keep Courts Fair, Impartial, and Independent*, 42 CT. REV. 22, 24 (2006); *see also* Pozen, *supra* note 55, at 272.

224. Brandenburg, *supra* note 223, at 24; Pozen, *supra* note 55, at 272.

225. *See supra* Part IV.

226. *See* Charles Gardner Geyh, *Why Judicial Elections Stink*, 64 OHIO ST. L.J. 43, 54-55 (2003).

The development and maintenance of transparently accountable JPE programs can help build this public trust as well as the trust of the judiciary and policy makers. The potential impact of this is impressive. With this increased ability to hold judges accountable, the public is likely to forestall efforts to curtail judicial independence at the ballot box. With increased ability to hold judges accountable, voters may vote to replace poor performing judges on the bench with greater frequency and, in the process, help increase the quality of the judiciary. With increased reliable information to use in judicial elections more voters will likely participate in selecting their judges. With increased trust in the method upon which they are evaluated judges may publicly support JPE programs and encourage voters to make informed decisions when voting in judicial elections. With reliable information about the effectiveness of JPE programs achieving their stated goals, policy makers and legislators are more likely to support their continued operation, and potentially their expansion.

Much can be gained by taking the steps needed to maintain effective and trustworthy JPE programs. It is up to those conducting the evaluations of the judges to evaluate themselves with an eye towards self-improvement and the betterment of the judicial system.