

1-15-1981

Evaluation of Judicial Performance: A Tool for Self-Improvement

Richard L. Aynes

Follow this and additional works at: <https://digitalcommons.pepperdine.edu/plr>



Part of the [Judges Commons](#), [Legal Ethics and Professional Responsibility Commons](#), and the [Legal Profession Commons](#)

Recommended Citation

Richard L. Aynes *Evaluation of Judicial Performance: A Tool for Self-Improvement*, 8 Pepp. L. Rev. Iss. 2 (1981)

Available at: <https://digitalcommons.pepperdine.edu/plr/vol8/iss2/1>

This Article is brought to you for free and open access by the Caruso School of Law at Pepperdine Digital Commons. It has been accepted for inclusion in Pepperdine Law Review by an authorized editor of Pepperdine Digital Commons. For more information, please contact bailey.berry@pepperdine.edu.

Evaluation of Judicial Performance: A Tool for Self-Improvement

RICHARD L. AYNES*

The quality of our judicial system, like other institutions, is a function of the work performed by those who are afforded major roles in the dispensation of justice. Unmistakably, judges, jurors and lawyers assume key roles in this process. Professor Aynes, who is a member of the A.B.A.'s Evaluation of Judicial Performance Committee, recognizes that both judges and lawyers, unlike jurors, are professionals expected to bring more to the bench than honesty, good faith and diligence. The author observes that while efforts to improve the daily performance of attorneys have been well under way since the early 1970's, it is now imperative that we address the issue of improving judicial performance. Professor Aynes reviews the historical antecedents of programs evaluating judicial performance, and shows the significant difference between the A.B.A.'s proposal and the various surveys and polls that have been previously utilized, while examining the critical issues confronting the profession in designing a responsible system for evaluating judicial performance.

I. INTRODUCTION

Under our system of government, the courts have traditionally been the arbitrators of many of the most important public and

* B.S. Miami University, Oxford, Ohio, 1971; J.D. Cleveland State University, 1974; Assistant Professor of Law, University of Akron. Mr. Aynes is currently serving as one of the Vice Chairpersons of the American Bar Association, Criminal Justice Section's Committee on the Evaluation of Judicial Performance. However, except where explicitly stated, the views expressed are his own and not necessarily those of the A.B.A., the Criminal Justice Section, or the Committee.

private disputes.¹ It is, therefore, natural to expect that the quality of justice is both a major concern and a high priority of our citizenry.² This concern may be expressed in a variety of ways as indicated by the results of polls of its membership by the American Bar Association³ and protests by citizens who believe that injustice occurred in a given case.⁴

Like other institutions, the quality of the judicial system turns upon the quality of the work performed by those involved in that system. Though others play key roles in the dispensation of justice,⁵ those who most affect its quality are jurors, lawyers and judges. Any serious attempt to make a significant improvement in the quality of justice must necessarily focus upon improving the performance of members of these three groups.

In considering the role that jurors play, it would seem that there is very little in the way of change which could improve the quality of justice. This statement is made not to deny the benefits and desirability of the various juror-oriented programs designed to educate jurors as to their function, and to make more efficient use of a juror's time.⁶ Rather, this conclusion is based upon the recognition that the basic contribution of the juror is "common sense"⁷ and his/her reflection of the conscience of the commu-

1. Scarcely any political question arises in the United States that is not resolved, sooner or later, into a judicial question. Hence, all parties are obliged to borrow, in their daily controversies, the ideas, and even the language, peculiar to judicial proceedings. As most public men are or have been legal practitioners, they introduce the customs and technicalities of their profession into the management of public affairs.

1 A. DE TOCQUEVILLE, *DEMOCRACY IN AMERICA* 290 (P. Bradley, ed., 1954).

2. Justice, sir, is the great interest of man on earth. It is the ligament which holds civilized beings and civilized nations together. Wherever her temple stands, and so long as it is duly honored, there is a foundation for social security, general happiness, and the improvement, and progress of our race.

2 W.W. STORY, *LIFE AND LETTERS OF JOSEPH STORY* 26 (1851) (quoting Daniel Webster).

3. In a recent Law Poll survey conducted by the American Bar Association, those lawyers responding indicated that the quality of the judiciary was their area of most concern for the profession. *What Are the Concerns of Lawyers?*, 66 A.B.A.J. 842 (1980).

4. *E.g.*, Goldstein, *Death Penalty Opponents Embittered by Executions*, N.Y. Times, May 26, 1979, at 6; Reynolds, *Blacks Countering KKK Terrorism*, Akron Beacon, J. Nov. 2, 1980, at col. 5.

5. Bailiffs, law clerks, court clerks, secretaries, deputies, and countless others are indispensable to a fair and efficient court system.

6. *E.g.*, BIRD ENGINEERING-RESEARCH ASSOCIATES, *LAW ENFORCEMENT ASSISTANCE ADMINISTRATION, A GUIDE TO JURY SYSTEM MANAGEMENT* (1975); LAW ENFORCEMENT ASSISTANCE ADMINISTRATION, *A GUIDE TO JUROR USAGE* (1974); and CARLSON, HALPER & WHITCOMB, *ONE DAY/ONE TRIAL JURY SYSTEM* (1977).

7. *See generally* *Duncan v. Louisiana*, 391 U.S. 145, 155-56 (1968); *State v. Payne*, 7 Ohio Op. 3d 432, 451-52 (Jackson, J., dissenting) (1978).

nity.⁸ As long as the juror brings to court an open mind, a capacity to listen to the evidence, a desire to search diligently for the truth, and a willingness to follow his/her oath, as juror, we can ask for nothing more.

The lawyer, unlike the juror, is an expert whose expertise, or lack thereof, can have a tremendous impact upon the outcome of a given case, despite good faith efforts to represent the client in the most professional manner possible. Since the lawyer's performance requires skill as well as judgment, there is tremendous potential for variation in the quality of representation provided by different lawyers.

Consequently, great efforts have been made to ensure that only qualified individuals become lawyers and to make provision for expulsion from the bar of those who are unsuited to continue. No longer can one become a lawyer after cursory study⁹ and an even more cursory oral examination at the bench.¹⁰ Requiring an un-

8. *Witherspoon v. Illinois*, 391 U.S. 510, 519 (1968).

9. Before Langdell, preparation for the bar was haphazard and lawyers with "marginal abilities were common and in all likelihood predominated." Devitt, *The Case Against the Case Method*, NAT'L. L.J., Apr. 21, 1980, at 15. See generally L. FRIEDMAN, *A HISTORY OF AMERICAN LAW* 278-82 (1973). While numerous other examples can be cited, the image of an ill-prepared Abraham Lincoln being admitted to the bar after cursory examination is the one that is most pronounced. To be faithful to the record, not all early requirements for bar passage were haphazard. Ohio, for example, required all applications to the bar to have "regularly and attentively studied the law, during the period of two years, previous to his application for admission." 1823 Ohio Laws 374. See also SWAN, *STATUTES OF THE STATE OF OHIO* 95 (1841). To balance Lincoln's experience in the study of law one may refer to his Illinois companion Richard Yates who, though never estimated as any more than a competent lawyer, nevertheless, was a "diligent" student whose study of the law was "so intense that in his twentieth year his eyes failed and he went totally blind for a short period of time." Northrup, *The Education of a Western Lawyer*, 12 AM. J. LEGAL HIST. 294, 297 (1968). George Sharswood, who became Justice of the Pennsylvania Supreme Court, described his years of reading law as ones of "assiduous employment and of earnest interrupted study." Heaney, *The Legal Education of George Sharswood*, 2 AM. J. LEGAL HIST. 259, 293 (1958). Sharswood's plan of study called for him to rise at four or five a.m. and included hours of study both before and after putting in an eight or nine hour day with his preceptor. Indeed, even Lincoln considered the supervision of an aspiring lawyer's education important enough that he refused to accept students in his office when he was unavailable to supervise them. E.g., 2 A. BEVERDIGE, *ABRAHAM LINCOLN* 142 (1928); (letter from Abraham Lincoln to James Thornton refusing to supervise J. Widner); Goff, *The Appointment, Tenure, and Removal of Territorial Judges: Arizona—A Case Study*, 12 AM. J. LEGAL HIST. 211, 215 (1968) (refusing to supervise I. Reavis).

10. In surveying the effect of the oral bar exams, L. Friedman had concluded that they were ineffective: "'examination' of lawyers-to-be in court was almost a farce, if we can believe a mass of anecdotes." L. FRIEDMAN, *A HISTORY OF AMERI-*

dergraduate degree¹¹ before the admission to law school; graduating from law school before taking the bar exam;¹² and passing a comprehensive written bar examination before being admitted into practice¹³ all increased the rigor of entrance into the bar. Further, the establishment of accrediting agencies for law schools¹⁴ tends to ensure that those academic requirements are not mere procedural hurdles. At the other end of the spectrum, the American Bar Association (A.B.A.) has taken the lead in formulating a code of professional conduct to regulate the ethical conduct of members of the bar¹⁵ and formulating standards for the discipline of those attorneys who become unfit.¹⁶

Yet, with the major exception of the publications of the A.B.A. standards,¹⁷ the profession had generally ignored the problems of

CAN LAW 277 (1973). Examples supporting this claim extend from the 1830's when Gustave Koerner, who emigrated from Germany and began practice in Illinois, reported his examination for the Illinois bar as being very informal, lasting only a half hour. *Id.* This practice continued through the latter part of the nineteenth century (William Howard Taft, 1880) Brockman, *The National Bar Association 1888-1893* 10 AM. J. LEGAL HIST. 122 (1966) and into the twentieth century. (Huey Long, 1915). T. WILLIAM, HUEY LONG 80-81 (1969).

On the other hand, their problems in obtaining admission to the bar had more to do with social status than knowledge. The difficulties of future Congressmen Joshua Giddings and Thaddeus Stevens had in obtaining admission to the bar indicates that the process was not as automatic as we have been led to believe. G. JULIAN, *THE LIFE OF JOSHUA R. GIDDINGS*, 23-24 (1892); BRODIE, *THADDEUS STEVENS* 32 (1966).

11. The accreditation standards require only three years of undergraduate education. ASSOCIATION OF AMERICAN LAW SCHOOLS, Association Information, BY-LAWS AND EXECUTIVE COMMITTEE REGULATIONS at 9, § 6-2, (1980); A.B.A., APPROVAL OF LAW SCHOOLS at 16, (1979). As a practical matter most admittees are required to have a bachelors degree. *See, e.g.*, NORTHWESTERN UNIVERSITY SCHOOL OF LAW, BULLETIN 59 (1979-81); PEPPERDINE UNIVERSITY SCHOOL OF LAW, BULLETIN 21 (1980-81).

12. *See generally* NATIONAL CONFERENCE OF BAR EXAMINERS, RULES FOR ADMISSION TO THE BAR IN THE UNITED STATES AND TERRITORIES (West, 1979).

13. *Id.*

14. ASSOCIATION OF AMERICAN LAW SCHOOLS, *supra* note 11, at 5-28 (1980); A.B.A., *supra* note 11.

15. A.B.A., CODE OF PROFESSIONAL RESPONSIBILITY AND CODE OF JUDICIAL CONDUCT (1978). *See also* A.B.A. DISCUSSION DRAFT OF ABA MODEL RULES OF PROFESSIONAL CONDUCT, 48 U.S.L.W. 1 (1980).

16. A.B.A. STANDARDS FOR LAWYER DISCIPLINARY AND DISABILITY PROCEEDINGS (tentative draft, 1978). *See e.g.* Ohio Sup. Ct. R. Gov't Bar 5(6).

17. The individual volumes published by the American Bar Association include STANDARDS RELATING TO PRETRIAL RELEASE (1968); PROVIDING DEFENSE SERVICES (1967); FAIR TRIAL AND FREE PRESS (1968); PLEAS OF GUILTY (1967); SPEEDY TRIAL (1968); JOINDER AND SEVERANCE (1968); SENTENCING AND ALTERNATIVES AND PROCEDURE (1968); APPELLATE REVIEW OF SENTENCES (1967); POST-CONVICTION REMEDIES (1967); DISCOVERY AND PROCEDURE BEFORE TRIAL (1970); PROBATION (1970); CRIMINAL APPEALS (1969); ELECTRONIC SURVEILLANCE (1970); THE PROSECUTION FUNCTION AND THE DEFENSE FUNCTION (1971); THE FUNCTION OF THE TRIAL JUDGE (1972); THE URBAN POLICE FUNCTION (1973); COURT ORGANIZATION (1973); and TRIAL COURTS (1976). The A.B.A., JUDICIAL DISCIPLINE AND DISABILITY

daily competence and performance.¹⁸ It was only in the early 1970's that the profession turned its attention to the question of helping lawyers improve their daily performance in the practice of law.¹⁹

Though the problems of daily performance have not been solved in their entirety, most law schools now offer a variety of skills courses designed to provide the students with a better foundation for future performance.²⁰ The A.B.A. has issued a special report emphasizing skills courses,²¹ and standards have been drafted for clinical legal education.²² The cooperative efforts of the A.B.A., the National Trial Lawyers Association and others, have resulted in post-law school education in trial practice.²³ Many states have now adopted mandatory continuing legal education.²⁴ Finally, the American Law Institute-American Bar Association Committee on Continuing Professional Education has developed "A Model Peer Review System."²⁵ This program, developed over a three-year period, includes model standards for

RETIREMENT was prepared by the A.B.A. Joint Commission on Professional Discipline and approved by the A.B.A. House of Delegates in February 1978.

18. The organized bar traditionally has preferred to deal with the problem of unethical practices while ignoring the issue of professional competence. Moreover, the professional grievance procedures are geared to ascertaining the culpability of practitioners. This preoccupation with fault can be said to limit, if not preclude, responsible criticism and corrective action by professional peers.

Dear, *Adversary Review: An Experiment in Performance Evaluation*, 57 DEN. L.J. 401, 405 (1980) (footnotes omitted).

19. Much of the credit for this change must be given to the Chief Justice of the Supreme Court. *E.g.*, Burger, *The Special Skills of Advocacy: Are Specialized Training and Certification of Advocates Essential to Our System of Justice?*, 42 FORDHAM L. REV. 227 (1973). One need not agree with the Chief Justice's assessment of the problem, either in its scope or proposed remedy, in order to acknowledge that his criticism sparked debate and change.

20. *E.g.*, NORTHWESTERN UNIVERSITY; *supra* note 11, at 14-15; PEPPERDINE UNIVERSITY, *supra* note 11, at 75-76; UNIVERSITY OF AKRON, SCHOOL OF LAW BULLETIN 14-15 (1979).

21. A.B.A. TASK FORCE ON LAWYER COMPETENCY, REPORT AND RECOMMENDATIONS ON THE ROLE OF THE LAW SCHOOLS (1979).

22. CLINICAL & LEGAL EDUCATION, REPORT OF THE A.A.L.S.-A.B.A. COMMITTEE ON GUIDELINES FOR CLINICAL EDUCATION (1980).

23. The National Institute for Trial Advocacy is such an effort. *See* HERMAN, MOSS, BROUN & SECKINGER, NITA TEACHERS MANUAL AND ADMINISTRATIVE GUIDE (3rd ed. 1979); J. SECKINGER & K. BROUN, PROBLEMS IN TRIAL ADVOCACY (1977); J. SECKINGER & K. BROUN, CASES IN TRIAL ADVOCACY (1977).

24. The states requiring continuing legal education include Colorado, Idaho, Iowa, Minnesota, North Dakota, South Carolina, Washington, Wisconsin, and Wyoming. OHIO LEGAL CENTER INSTITUTE, UPDATE 80 19.04 (1980).

25. *See* A.B.A., *Professional Competence*, 49 U.S.L.W. 2093 (1980).

evaluating attorney competence in eight separate areas.²⁶

Though these are pioneer efforts which will have to be studied and evaluated, the legal profession is on the road to improving the daily performance of the profession.

Unfortunately, the same cannot be said with respect to the judiciary. Judges, like lawyers and unlike jurors, are professionals who are expected to bring more to the bench than honesty, good faith, and diligence. The skill with which they perform their function has a tremendous effect upon the quality of justice. Because of this fact, there is a tremendous potential for improving the quality of justice by improving the performance of judges on the bench.

In many ways the profession's efforts at maximizing the quality of the judiciary today is analogous to its past efforts with respect to maximizing the quality of lawyer performance. Much time and effort has been spent to ensure that only qualified individuals become judges.²⁷ Similarly, judicial conduct may be guided by the Code of Judicial Conduct²⁸ and, in the event judges violate the trust given to them, they can be removed from office.²⁹ With this concentration upon obtaining quality people for judgeships and removing those who may later become unfit for such a position, the profession lacks any serious program designed to improve on the bench's judicial performance.³⁰

Since efforts to improve the daily performance of attorneys

26. Those categories include the following: (1) information gathering, (2) legal analysis, (3) strategy formation, (4) strategy execution, (5) practice management, (6) professional responsibility, (7) training and supervision, and (8) continuing attorney self-education. *Id.*

27. *E.g.*, L. BERKSON & S. CARBON, *THE UNITED STATES CIRCUIT JUDGE NOMINATING COMMISSION: ITS MEMBERS, PROCEDURES AND CANDIDATES* (1980); G. WINTERS, *JUDICIAL SELECTION AND TENURE* (1973); S. ESCOVITZ, F. KURLAND & N. GOLD, *JUDICIAL SELECTION AND TENURE* (1975).

28. A.B.A., *CODE OF PROFESSIONAL RESPONSIBILITY AND CODE OF JUDICIAL CONDUCT*, (1978).

29. W. BRAITHWAITH, *WHO JUDGES THE JUDGES* (1971) surveys the removal and retirement procedures for incompetence or disability in Missouri, New Jersey, New York, Illinois, and California. *See also A.B.A. Annual Meeting*, 49 U.S.L.W. 2125-26 (1980) which indicated that the House of Delegates had urged the adoption of legislation providing for the removal of unfit judges and for sanctions short of removal.

30. In taking this position, the author is fully aware and appreciative of significant contributions to improving judicial performance made by the National Judicial College, the Institute of Judicial Administration, N.Y.U. School of Law, the National Center for State Courts, and others. *E.g.*, *The National Judicial College 1980 Calendar of Resident Sessions*, JUDGES J., Spring, 1980, at back cover; Cameron, *Second Degree for Appellate Judges*, 19 JUDGES J., Spring, at 35. As discussed more fully below, the proposed evaluation of judicial performance would complement those existing programs. *See also* Handler, *A New Approach to Judicial Evaluation to Achieve Better Judicial Performance*, St. Ct. J., Summer, 1979, at 7.

seem to be well under way, it is only logical to now address the issue of improving judicial performance. The benefits of such a program could have a large effect because the average judicial career is twenty-five years in length³¹ and because this is a time of expansion of the judiciary in both the state³² and federal systems.³³

Like many ideas "whose time has come," the use of evaluations of judicial performance to improve the quality of the work of judges who are already on the bench is an idea which probably came to many people across the country almost simultaneously. It is currently being pursued not only by the A.B.A., but also by the courts in New Jersey, Colorado, and other concerned organizations including the National Center for State Courts. The purpose of this article is to review the origins of the proposal for a program to evaluate judicial performance, show the significant difference between the A.B.A.'s proposal for evaluating judicial performance and the various surveys and polls that have been utilized in the past, and to explore the issues in establishing a responsible system for evaluating the performance of judges.

31. G. WINTERS, JUDICIAL SELECTION AND TENURE 212-13 (1973).

32. *E.g.*, OHIO REV. CODE ANN. § 1901.01 (Baldwin 1975) (creating new municipal courts); OHIO REV. CODE ANN. §§ 2301.02-2301.03 (Baldwin 1978) (creating additional common pleas judgeships).

33. The Omnibus Judgeship Bill, 28 U.S.C. §§ 44, 133 (Supp. 1980), created 152 new judgeships. While a growing caseload has generally been thought to threaten the quality of justice, and increasing the number of judges has generally been seen as an answer to that threat, *e.g.*, Rosenn, *Trends in Administration of Justice in the Federal Courts*, 39 OHIO ST. L.J. 791, 794-96 (1978); there may also be disadvantages to that solution. For example, consider the difference between a president appointing ten judges from a given pool of 1000 attorneys and appointing 100 judges from that same pool. Conventional wisdom would suggest that while all the appointees might be qualified in both situations, the appointment of the smaller number of judges would probably result in a higher caliber judiciary. Thus, by the very expansion of the number of judicial positions, one takes the chance that the average ability of the appointees will decline.

Similarly, adding new judges to existing courts changes the nature of the institution beyond what mere numbers might suggest. New judges mean new quarters, new secretaries, new law clerks, and a host of related employee, space, and bureaucratic problems. "Over the long run, augmenting judicial capacity may erode the distinctive contribution the courts make to the social order. The danger is that courts, in developing a capacity to improve on the work of other institutions, may become altogether too much like them." Shattuck & Norgren, *Political Use of the Legal Process by Black and American Indian Minorities*, 22 HOW L.J. 1, 23 (1979) (citing A. Horowitz, *THE COURTS AND SOCIAL POLICY* 298 (1977)). This being the case, new efforts may be required simply to maintain the current quality of judicial performance.

II. ORIGINS OF THE PROPOSAL

Job performance evaluation is an accepted management tool utilized in private industry,³⁴ the public sector,³⁵ and education.³⁶ It operates on a very basic premise of improving human behavior: that desirable conduct should receive positive reinforcement and that areas for improvement should be identified so that the individual in question can improve performance. As set forth below,³⁷ the current surveys and polls evaluating judicial candidates do not really attempt to assess on-the-job performance in a manner which can lead to self-improvement.³⁸ It is this void that the proposal for evaluating judicial performance seeks to fill. Such a proposal needs to provide a comprehensive, reliable method of assessing the performance of judges at their jobs. This assessment of performance needs to be sensitive and responsible so that it will not detract from judicial independence or performance of the traditional judicial duties. Finally, such a proposal needs to provide a basis for members of the bench to maximize their potential for excellence through self-improvement.

The impetus for such a program of judicial evaluation is difficult to trace. In New Jersey, the State Supreme Court established a special Committee on Judicial Evaluation and Performance, and that committee had submitted its report calling for such a program in early March of 1978.³⁹ The National Center on State Courts had also given the matter its attention and by September of 1979, developed a paper for internal use which called for fund-

34. NAT'L. CENTER FOR STATE COURTS, JUDICIAL PERFORMANCE EVALUATION, CONCEPT PAPER (draft, 1979).

35. *E.g.*, Jensen, *Employee Evaluation*, GRANTSMANSHIP CENTER NEWS, July-August, 1980, at 36.

36. Armour, *What Do They Expect of Me?*, CHRONICLE HIGHER EDUC., Oct. 15, 1979, at 48.

Every professor at the Georgetown University Law Center receives a rigorous appraisal by the students. They fill out highly developed and sophisticated forms covering almost 30 items regarding professorial performance. These cards are fed into a computer and the results are made available to the students and professors. The operation is run by students and is accepted by the Law School community. With rare exception the students assess professors in a responsible manner and I have personally benefited from their evaluation of me. No less could be expected of lawyers.

Letter from Herbert S. Miller to Tom Karas (February 15, 1979), at 2.

37. See text accompanying notes 56-66 *supra*.

38. This observation is not meant to denigrate the value of those efforts. To the contrary, it simply recognizes that because those instruments focus upon selection, retention, promotion and consideration, they cannot be expected to lend themselves to specific improvements in judicial performance.

39. Handler, *A New Approach to Judicial Evaluation to Achieve Better Judicial Performance*, ST. CT. J., Summer, 1979, at 3.

ing of a demonstration project.⁴⁰ Within the A.B.A., the concept was apparently first raised at the fall Council Meeting in 1978⁴¹ by Herbert S. Miller, then Co-Director of Georgetown University Institute of Criminal Law & Procedure, and currently Chairperson of the Criminal Justice Section of the A.B.A.

A.B.A. consideration of the concept of evaluating the performance of sitting judges was given expeditious treatment. In the summer of 1979 Chairperson Richard Gerstein of the Criminal Justice Section established the Evaluation of Judicial Performance Committee (hereinafter Evaluation Committee).⁴² Through the Evaluation Committee's Chairperson, Richard Kuh, and the A.B.A. Criminal Justice Section Director, Laurie Robinson, members of the Evaluation Committee had an opportunity in the fall of that year to review the existing literature on judicial evaluation.⁴³

The first meeting of the Evaluation Committee was held in Washington, D.C. on October 12 and 13, 1979. In addition to the diversity of the committee itself, which consisted of members of the state and federal bench, members of the bar in public and private practice, and three law faculty members, the Evaluation Committee benefited greatly by the presence of two members of the staff from the National Center for State Courts. They also had the opportunity to review the draft of the paper which that organization had prepared on evaluating judicial performance. After

40. NAT'L. CENTER FOR STATE COURTS, *supra* note 34.

41. See generally Letter from Herbert S. Miller, *supra* note 36.

42. The Evaluation Committee was comprised of the following persons: Richard H. Kuh, a New York City private practitioner, Chairperson; Federal District Court Judge Charles R. Richey of Washington, D.C., Vice-Chairperson; Boston criminal defense lawyer, Paul T. Smith, Vice-Chairperson; Washington, D.C. attorney Herbert S. Miller, Chairperson-Elect of the Criminal Justice Section; Boston criminal defense lawyer Joseph Balliro; Linda Cole of the Justice Department Civil Division; Judge Burton Katz of the Los Angeles Municipal Court; Richard L. Aynes, Assistant Professor of Law, University of Akron School of Law; Tampa, Florida lawyer Paul B. Johnson, former Florida State's Attorney; David Horowitz of the Los Angeles Public Defenders Office; attorney Candace Kovacic of Washington, D.C.; Jane Hazen, a Denver criminal defense attorney; Judge Lenore Nesbitt of Miami's Circuit Court; Justice Rosalie Wahl of the Minnesota Supreme Court; Dean Emeritus Robert B. Yegge of the University of Denver Law School; Professor Douglas Haddoc of Hamlin University School of Law in St. Paul, Minnesota; Justice William G. Callow of the Wisconsin Supreme Court; and Seattle attorney Morris H. Rosenberg.

43. Of particular value was Justice Handler's article concerning the Report of the New Jersey Supreme Court's Committee on Evaluation and Performance, Handler, *supra* note 39.

two days of debate and discussion, the Evaluation Committee determined that the goal of effective judicial performance evaluation could best be pursued by a special project committee composed not only of representatives of the various A.B.A. entities and affiliates, but also of representatives of other major organizations having an interest in such a project. The Evaluation Committee urged the A.B.A. to sponsor this cooperative effort.⁴⁴

The Council of the A.B.A.'s Criminal Justice Section endorsed the Evaluation Committee's work by unanimously adopting a resolution at its November 1979 meeting calling on the A.B.A. leadership to "name a special committee or task force (or such other means as the leadership deems appropriate) to propose standards for the evaluation of judicial performance."⁴⁵ Following the

44. The Evaluation Committee reached the following consensus:

1. That judicial evaluation was an idea whose "time had come"; that it was going to be done by someone; and that the legal profession should take a leadership role to ensure that it was done in a sensitive and professional manner which would be useful and, at the same time, not infringe upon the necessary independence of the judiciary.

2. That the issue of judicial evaluation clearly transcended the jurisdiction of the Criminal Justice Section and that it was important, as quickly as possible, for a special project committee to be established that would include representation from all interested divisions of the A.B.A. and affiliated organizations, as well as other groups that had a vital interest in such a project, including the National Center for State Courts, the American Judicature Society, the National Judicial College, and other similar organizations.

3. That the issues involved in pursuing development and implementation of a meaningful program of judicial evaluation were extremely complex and were not capable of resolution by our committee given its limited resources and expertise. Further, even if the committee could resolve those issues, it should not do so: they were ones upon which reasonable people can disagree and should be resolved only after plenary consideration of the views of all interested parties through the vehicle of a special project committee.

4. That while the committee was not in a position to resolve these issues, it nevertheless had a responsibility to attempt to bring the importance of this program to the attention of others and to attempt to secure a broad base of support for the investigation of the desirability and feasibility of establishing a program of judicial-performance evaluation.

5. In order to meet its responsibility of presenting this idea to other entities both within and without the A.B.A., the Committee determined to draft a Concept Paper of its own which would indicate the recommendation of the Committee that some program in evaluating judicial performance should be established and attempt to indicate the issues which a special committee would have to resolve before such a program could take final form for consideration and implementation. In attempting to outline the "issues," or problems, involved in judicial evaluation, the Committee agreed that the objective was to attempt to demonstrate the seriousness and complexity of the undertaking and not to predetermine the resolution of those questions.

See Resolution of the A.B.A. Evaluation of Judicial Performance Committee (October 13, 1979).

45. Memorandum from Laurie Robinson, Section Director, A.B.A. Criminal Justice Section, to Members of the Evaluation Committee (November 15, 1979).

course of action agreed upon at its Washington, D.C. meeting,⁴⁶ the Evaluation Committee drafted a Concept Paper⁴⁷ setting forth the reasons why there was a need for a comprehensive inquiry into the question of judicial performance.⁴⁸ The Council's resolution and the Evaluation Committee's Concept Paper were distributed in mid-January of 1980, to the A.B.A. committees and to other groups who have a common interest in this area.⁴⁹ This allowed the February Mid-Year Meeting of the A.B.A. in Chicago to serve as the initial forum for public discussion of the proposal.⁵⁰

46. The Evaluation Committee decided to divide the task of preparing a Concept Paper among three subcommittees: a criteria-indicators subcommittee to explore the issues involved in setting the standards by which judges should be evaluated; a methodology subcommittee to explore the difficulties in seeking to obtain information necessary to evaluate judges; and a uses-purposes subcommittee to explore the alternatives to which the data and evaluations could be put after the evaluations had been conducted.

After the Washington meeting, a bibliography and additional materials were distributed to all members of the Evaluation Committee. Further ideas and information were exchanged between members of the various subcommittees, as well as between the section staff and members of the National Center for State Courts. By the end of December, the work of the subcommittees had been completed.

The subcommittee drafting the section on criteria and indicators was chaired by Dean Yegge. The subcommittee drafting the section on uses was chaired by Justice Wahl. The subcommittee drafting the section on methodology was chaired by the author of this article.

47. A.B.A. EVALUATION OF JUDICIAL PERFORMANCE COMMITTEE, CONCEPT PAPER (1979) [hereinafter cited as CONCEPT PAPER].

48. This Concept Paper in general, and the Methodology Section in particular, are the foundation upon which much of the latter portion of this article is based.

49. The organizations which received this information included the following A.B.A. sections: The Section of Antitrust Law, Section of Corporation, Banking and Business Law, Section of Insurance, Negligence and Compensation Law, Section of General Practice, Section of Individual Rights and Responsibilities, Judicial Administration Division, Appellate Judges Conference, National Conference of State Trial Judges, National Conference of Special Court Judges, National Conference of Federal Trial Judges, Law Student Division, Section of Litigation, Section of Taxation, Young Lawyers Division, Committee on Advanced Judicial and Legal Education, Committee on Association Standards for Criminal Justice, Committee on Ethics and Professional Responsibility, Committee on Coordination of Federal Judicial Improvements, Committee on Federal Judiciary, Committee on Judicial Selection, Tenure and Compensation, and the Committee on Professional Discipline.

The following organizations also received the Concept Paper and the Evaluation Committee's resolution: American Judicature Society, American Law Institute, Conference of Chief Justices, National Conference of Commissioners on Uniform State Laws, National Judicial College, American Bar Foundation, and the National Center for State Courts. Memorandum from Richard H. Kuh, to members of the Evaluation of Judicial Performance Committee (January 15, 1980).

50. See, e.g., Burke, *Bar Unit Urges National Rating of Judges*, NAT'L L.J., February 18, 1980, at 9; HARWELL, A GUIDE TO THE NORTH CAROLINA JUDICIARY 6 (1979).

Other organizations concerned about the quality of justice were quick to indicate their interest in pursuing an inquiry into judicial performance evaluation. At this time support for such an inquiry includes many sections within the A.B.A. including the Anti-Trust Section, Insurance Section, Association Standards for Criminal Justice Committee, and the Professional Discipline Committee. Dean Watts of the National Judicial College, the National Center for State Courts,⁵¹ the National Conference of Special Trial Judges and the Judicial Administration Division Council have indicated their support for the inquiry as well.⁵²

As an initial step toward this broad-based inquiry into judicial performance, the A.B.A. Judicial Administration Division, the Section of Litigation, and the Section of Criminal Justice jointly sponsored a panel discussion of the evaluation of judicial performance at the A.B.A. annual meeting in Honolulu in August, 1980.⁵³ Former A.B.A. President, Leonard S. Janofsky designated this as a "special emphasis" program.

At the annual meeting, the Board of Governors of the A.B.A. authorized the A.B.A. president to appoint a broadly representative seven-member Special Committee on the Evaluation of Judicial Performance to investigate the establishment of standards and to establish a demonstration project on implementing such standards for a program of judicial review. Simultaneously, authorization was given to seek funding for the project.⁵⁴ Such a funding

See also Standard-Setting Urged for Judicial Performance, CRIM. JUST., May, 1980, at 4.

51. Memorandum, from Richard H. Kuh, to the members of Evaluation of Judicial Performance Committee (June 6, 1980). As expected, not all of the responses were favorable. The A.B.A. Young Lawyers Division indicated its commitment to a lawyer-polling process and its feeling that the project was of limited usefulness. The A.B.A. Judicial Selection, Tenure and Compensation Committee indicated that it did not support the project, because that committee felt it was duplicative of D. MADDI, JUDICIAL PERFORMANCE POLLS (1977). The American Judicature Society indicated that it believed that such a standard-setting project was premature.

52. Memorandum from Richard H. Kuh, to members of the Evaluation of Judicial Performance Committee (October 15, 1980).

53. Richard H. Kuh moderated the panel which included Chief Justice Ray Rabinowitz of the Alaska Supreme Court; Dorothy Maddi of the American Bar Foundation and author of the significant study *Judicial Performance Polls*; Austin Jennings, a prominent Arkansas trial lawyer; and Scott Armstrong, co-author of *The Brethren*. See *Evaluating Judges: A "Thorny Thicket"*, 66 A.B.A.J. 1052 (1980).

54. Memorandum, from Richard H. Kuh, to members of the Judicial Performance Committee (September 2, 1980). The A.B.A. president was also given authorization, in the alternative, to work with the Criminal Justice Section in seeking an existing A.B.A. entity to oversee the project. A.B.A. President William Reece Smith, Jr., has now determined to utilize the first option, of appointing the special committee once funding is secured. Note 52 *supra*. This is consistent with the strongly held feelings of the members of the Evaluation of Judicial Performance Committee that the success of any such project depends upon a broadly based

proposal, prepared by the Director of the Criminal Justice Section, Laurie Robinson, is pending before the Law Enforcement Assistance Administration.⁵⁵

Thus, the A.B.A. is at the threshold of establishing a broad-based inquiry leading to the establishment of standards and pilot programs in evaluating judicial performance. Because the prospects for such a project seem promising, it is appropriate to consider the issues involved in pursuing such an inquiry, such considerations should serve a number of purposes including;

1. Setting forth with more precision the goals of the judicial performance evaluation;
2. Indicating the necessity for broad-based participation on the Special Committee;
3. Demonstrating the complexity of many of the issues and why adequate funding is so crucial to the success of the project; and
4. Bringing these questions to the attention of other interested persons who may begin to think about these problems and may generate additional ideas which have not been considered by those who have been working on this project in the past.

With this in mind, consideration should be given to the following questions: how will the work of the proposed seven-person committee differ from past efforts; what criteria should be used in evaluating judges' performance; what methodology should be utilized in gathering data for the evaluation; and to what uses should the data be put?

III. THE DIFFERENCE BETWEEN BAR POLLS AND EVALUATING JUDICIAL PERFORMANCE

Two initial hurdles will have to be overcome by the A.B.A. Special Committee on the Evaluation of Judicial Performance. The first is the perception that this proposal to evaluate judicial performance is simply a call for an additional poll of attorneys by the bar association to vote upon whether or not judges are performing

oversight committee, drawing upon the different views of various entities both within and without the A.B.A.

55. The funding proposal calls for the establishment of a seven-person standing committee, appointed by the A.B.A. president from a cross section of interested entities to supervise the project. In addition to normal support staff, the proposal calls for the creation of two five-person task forces, one to consider criteria or indicators for judicial evaluation and the other to consider the uses, purposes and methodology problems. The drafting of standards themselves is projected on a 16-month timetable, with required funding of approximately \$200,000. The implementation stage would involve cooperative efforts with other groups in education and the development of pilot projects in several jurisdictions.

their jobs properly.⁵⁶ The second hurdle is the perception that all of the necessary study upon the issues involved in such polling was published by the American Bar Foundation in 1977.⁵⁷ It is submitted that neither perception is correct.

A. Polling is an Inadequate Substitute for Systematic Evaluation

Traditionally, judges have been evaluated, if at all, by surveys or polls of attorneys which are usually conducted by local bar associations. Though the format obviously varies from poll to poll, the end result is usually to elicit from members of the bar their subjective opinions about the judges' overall performance, as well as their opinions of a number of general traits such as honesty, fairness and lack of bias. As a general rule the results of such polls are utilized by third parties, such as voters or public officials, in making choices about retaining or promoting the judges who were the subjects of the poll. In contrast, as used in this article, the proposal to evaluate judicial performance refers to a systematic, multi-faceted attempt to gather data, subjective and objective, which would give judges insight into their performance in such a manner as to reinforce that performance when it is desirable and to spur improvement where improvement could be obtained.

There is a significant difference between traditional polling and this proposal to consider evaluating judicial performance. There has been a wealth of bar polls and some fine studies of their methodology and impact.⁵⁸ This experience is obviously beneficial and offers much to build upon when considering judicial performance evaluation. At the same time, it must be noted that the objectives of these polls were different from a proposal to improve on-the-bench performance.

First, in spite of the use of the term "evaluation of judicial performance", in most instances these efforts were aimed at selection or retention of a judge.⁵⁹ As such, those polls were little more

56. *E.g.*, Letter from David C. Weiner, Chairman, A.B.A. Young Lawyer's Division, to Richard H. Kuh (April 15, 1980).

57. D. MADDI, *supra* note 51.

58. Three such studies were of particular importance to the Evaluation Committee: D. MADDI, *supra* note 51; J. GUTERMAN and E. MEIDINGER, AMERICAN JUDICATURE SOCIETY, IN THE OPINION OF THE BAR (1977); C. PHILIP, INSTITUTE OF JUDICIAL ADMINISTRATION, HOW BAR ASSOCIATIONS EVALUATE SITTING JUDGES (1976).

59. In fact, at the present time the only program specifically aimed at improving judicial performance by evaluation, as opposed to election or retention, is that proposed in New Jersey. See generally Handler, *supra* note 30, at 3. Apparently Colorado is currently considering a similar program. Letter from E. Keith Stott,

than straw votes which could have almost no impact upon improving the performance of a particular judge. Without being specifically designed for on-the-bench improvement, the educational benefit of such polls is necessarily going to be only incidental and sporadic. Where the goal is not related to retention or selection, but rather toward improvement of sitting judges, a totally different methodology may be necessary.⁶⁰

Second, because the polls are often keyed to retention or selection, they, like the grievance procedures filed against attorneys, create threats and resistance to a fair consideration for judicial improvement. It is hard for anyone, whether it be an attorney looking at a grievance or a judge looking at an unfavorable bar poll, to accept the claim that this is mere constructive criticism when so much is at stake.

Third, there is widespread criticism of the existing polls and polling procedures which casts doubt upon their reliability. Such criticism generally keys in on the low response rate by those being asked to complete the poll⁶¹ and includes the point that even the information provided by those who do respond may be unreliable. To buttress the latter claim, examples are brought forth where judges who had not been on the bench for a number of years were accidentally placed in the poll and received high ratings.⁶²

Further, the judges themselves seem to be particularly critical of such bar polls. One judge, as a member of the Evaluation of Judicial Performance Committee, reviewed numerous polls used in several states and "found all of them to be wanting".⁶³ A federal judge once so strongly disagreed with a bar poll, that he

Jr., Colorado Deputy State Court Administrator, to Los Angeles Municipal Court Judge Burton Katz (November 8, 1979).

60. The utilization of straw polls is indicative of a selection, rather than an evaluation orientation on the part of the administering authority. In themselves, the results from a straw poll do not provide a sitting judge with any meaningful feedback on the basis of which he or she might modify his or her behavior, for they may reflect the strength and weaknesses of the non-incumbent opponent(s) as much or more than the judge's own.

J. GUTERMAN & E. MEIDINGER, *supra* note 58, at 36.

61. Marks, *Verdicts on Judges*, JURIS DOCTOR, Jan. 1977, at 39.

62. *Id.* at 40, contains a reference to a judge who had not served for eight years prior to the poll but was nevertheless included. *See also* Hook & Sparks, *Evaluating Judicial Evaluation*, BRIEF CASE, August, 1980, at 18-19.

63. Letter from Los Angeles Municipal Court Judge Burton Katz to Minnesota Supreme Court Justice Rosalie E. Wahl (November 1, 1979). Judge Katz found the polls to be inadequate.

wrote to object prior to the time it was taken.⁶⁴ A Los Angeles County judicial candidate who received a top rating took exception to the evaluation process.⁶⁵ These views receive support from independent analysis which concludes that "[e]ven the most meticulous polls to date have fallen so far short of a reasonable scientific standard" that the researcher advocated a "more journalistic" approach to rating judges.⁶⁶

One need not enter the debate as to the reliability or lack thereof of the existing polls to see their principal defect in attempting to improve judicial performance. They are perceived by many to be unreliable. It is unrealistic to believe, even if bar polls were specific enough to allow for self-assessment and self-improvement, that judges are going to change their conduct based upon a poll which they believe defective and unreliable.

For each of the foregoing reasons, it is submitted that bar polls cannot be equated with a systematic evaluation of judicial performance. While the polls may have their place in selection or retention, and may constitute a key element in a more comprehensive evaluation program, polls cannot take the place of a multi-faceted evaluation of judicial performance.

B. Further Study is Needed on Evaluating Judicial Performance

The related misconception is that the work of the proposed Special Committee on Evaluating Judicial Performance would be duplicative of Dorothy Maddi's *Judicial Performance Polls* published by the American Bar Foundation in 1977. Ms. Maddi's work was one of the pioneering efforts in the field and certainly has made a significant and lasting contribution which has been uniformly acknowledged. Rather than being viewed as the final work in the area, *Judicial Performance Polls* should be seen as the foundation upon which to build a much broader inquiry. Per-

64. Judge Andrew Hauk wrote that a poll to be conducted by the Barristers of Beverly Hills called for

a childish 'numbers game' of hearsay opinions by anonymous pollsters who in the vast majority have not practiced in any of our courts, cannot judge the judges upon anything other than what they have heard from other faceless anonymities, a good many of whom may well be disgruntled if experienced, and uninformed themselves if not experienced.

Marks, *supra* note 61, at 38.

65. They put you in a docket, like a defendant, but instead of giving you a bill of particulars, they hand you a list of generalities. It's very difficult to answer nebulous criticism, like, 'We hear you've been rude or abrupt with attorneys.'—An unidentified judgeship candidate who received a top rating in the Los Angeles County Bar Association evaluation process.

Evaluating Judges: A "Thorny Thicket", *supra* note 53, at 833.

66. Flanders, *Evaluating Judges: How Should the Bar Do It?*, JUDICATURE, Feb., 1978, at 304-05.

haps the best demonstration of this point is to be found in comparing *Judicial Performance Polls* with the Concept Paper developed by the A.B.A. Evaluation of Judicial Performance Committee.

Timeliness. *Judicial Performance Polls* was published in 1977. Most of the data relied upon was based upon a 1975 survey.⁶⁷ Thus, the information upon which the study is founded is somewhere between three and five years old. In addition to the normal expectation that numerous changes would have been made by the various organizations conducting judicial polls,⁶⁸ recently significant new proposals have been made for more sophisticated programs of judicial evaluation than existed in 1977.⁶⁹

Scope of Consideration. *Judicial Performance Polls* was intended to be a "study of how judicial performance polls *have been conducted*."⁷⁰ Though the purpose of such a study was to "develop guidelines for designing and conducting future polls,"⁷¹ such guidelines were, at least implicitly, to be a deduction from past practices. While the Concept Paper of the Evaluation Committee gave careful consideration to past practices,⁷² it also attempted to set forth alternatives that may not have been attempted before.⁷³ Thus, the scope of inquiry of the Concept Paper is considerably broader than that of *Judicial Performance Polls*.

Retention and Selection Objectives. *Judicial Performance Polls* is careful to make the important distinction between polls which

67. D. MADDI, *supra* note 51, at 2. See generally Appendix containing sample polling instruments; D. MADDI, *supra* note 51, at 27-107.

68. For example, two of the polls reviewed in the preparation of *Judicial Performance Polls* were those conducted by the Akron (Ohio) Bar Association and the Cincinnati (Ohio) Bar Association. D. MADDI, *supra* note 51, at 27. Since the publication of *Judicial Performance Polls*, both associations have changed their polling methods. In fact, rather than maintaining specificity as suggested by *Judicial Performance Polls*, D. MADDI, *supra* note 51, at 18-20, 24-25, both of those organizations have gone to less specific and more generalized criteria. See *Newspapers Missed the Real Story on Judicial Selection*, CINCINNATI BAR ASSOCIATION REPORT, Oct., 1979, at 5.

69. E.g., Handler, *A New Approach to Judicial Evaluation to Achieve Better Judicial Performance*, ST. CT. J., Summer, 1979, at 3; Letter from E. Keith Stott, Jr., *supra* note 59.

70. D. MADDI, *supra* note 51, at 1 (emphasis added).

71. *Id.*

72. See notes 27-50 *supra* and accompanying text.

73. CONCEPT PAPER, *supra* note 47, at 4. "In attempting to consider all possibilities for 'sponsors' of such an evaluation. . . ." *Id.* at 6: "In attempting to consider a number of reasonable sources, the following list is offered. . . ."

are merely plebiscites and those which actually attempt to evaluate judicial performance.⁷⁴ At the same time, *Judicial Performance Polls* indicates that of those polls studied, the attempts at improving judicial performance was targeted to three groups: (1) the electorate; (2) appointing officials; or (3) the judges being evaluated. The theory underlying the attempt to use poll results to influence the electorate and appointing officials was that the way to improve judicial performance is to increase the quality of the judges who are elected or appointed to the bench.

The primary focus of *Judicial Performance Polls* is upon the selection and retention process. In marked contrast, the Concept Paper focuses exclusively upon evaluations for the purpose of improving the performance of sitting judges and specifically eschews any attempt to use the evaluations for selection or retention purposes.⁷⁵ This is a particularly important distinction "[s]ince the articulated goal here is not related to retention or selection, but rather toward improvement of sitting judges, a *different methodology may be necessary*."⁷⁶

Purposes and Objectives. *Judicial Performance Polls* devotes one and a half pages to a summary of the reason the existing polls were conducted.⁷⁷ Though some question is raised about the effectiveness of such polls on judges who are not facing a retention or selection decision,⁷⁸ the discussion by *Judicial Performance Polls* is generally descriptive rather than analytical or exploratory. In the section discussing guidelines, no reference is made to purposes or objectives.⁷⁹ In contrast, the Concept Paper calls for plenary consideration of possible purposes of such an evaluation.⁸⁰ As starting points for discussion, the Concept Paper suggests consideration be given to the possible use of such evaluations for self-improvement, judicial education programs, and allocation and assignment of judges.

Sponsoring and Conducting Evaluations. *Judicial Performance Polls* has one half page devoted to a description of those who conducted those polls which were examined.⁸¹ The guidelines section of *Judicial Performance Polls* does not address the question of who should be involved in either the sponsoring or administra-

74. D. MADDI, *supra* note 51, at 1-2.

75. CONCEPT PAPER, *supra* note 47, at 2. "such evaluation is aimed at improving on-the-bench performance, rather than for use in selection or retention of judges."

76. *Id.* at 4 (emphasis added).

77. D. MADDI, *supra* note 51, at 3-4.

78. *Id.* at 4.

79. *Id.* at 15-16.

80. CONCEPT PAPER, *supra* note 47, at 9-11.

81. D. MADDI, *supra* note 51, at 3.

tion of the evaluations. In contrast, the Concept Paper calls for plenary consideration of the issue.

Sources of Information. Under the topic "Who Was Eligible to Participate in the Polls", *Judicial Performance Polls* addresses the critical issue of the attempt to insure a representative sample in surveying attorneys.⁸² *Judicial Performance Polls* assumed that attorneys, and only attorneys, are going to be the source of information for purposes of the poll or evaluation. This is equally true in the *Judicial Performance Polls'* section on guidelines.⁸³ The Concept Paper, in attempting to keep all alternatives open for future study, has suggested that all potential sources of evaluating information be explored and, has suggested no fewer than sixteen categories of individuals who might be considered.⁸⁴

Criteria or Indicators. The discussion in *Judicial Performance Polls* was limited to consideration of whether to use general, as opposed to specific, criteria for evaluation. *Judicial Performance Polls* considered the effect the number of those criteria would have upon the time it took to complete the forms.⁸⁵ Though the extensive appendix includes useful examples of types of criteria, *Judicial Performance Polls* does not attempt to evaluate these criteria or to suggest what criteria should be utilized in evaluating the performance of judges. On the other hand, the Concept Paper, in an ambitious undertaking, has called for specific standards to evaluate judicial performance and encourage improvements.⁸⁶

Dissemination of Results. *Judicial Performance Polls* summarizes the bar polls it considered⁸⁷ and, concluded that in order to maximize the influence of the poll, "full public disclosure is required."⁸⁸ The Concept Paper found that disclosure to the public was an open question which deserved careful study,⁸⁹ perhaps be-

82. *Id.* at 5-8.

83. *Id.* at 15-18.

84. CONCEPT PAPER, *supra* note 47, at 6-7. These categories are as follows: (1) The judges under evaluation; (2) other judges; (3) attorneys; (4) jurors; (5) witnesses; (6) parties; (7) court observers; (8) police officers; (9) probation officers; (10) courtroom spectators; (11) media representatives; (12) representatives of citizens groups; (13) law professors and other academicians; (14) court administrators; (15) law clerks and bailiffs; and (16) secretaries and clerks.

85. D. MADDI, *supra* note 51, at 9-11, 18-21.

86. CONCEPT PAPER, *supra* note 47, at 12-13.

87. D. MADDI, *supra* note 51, at 12-14. Twenty-two of the 32 polls considered were, in whole or in part, released to the public.

88. *Id.* at 22.

89. CONCEPT PAPER, *supra* note 47, at 11.

cause of the feeling that public disclosure was more closely related to retention or selection decisions than to the improvement of the performance of sitting judges.

Methodology. Though *Judicial Performance Polls* discusses some significant problems involved in the methodology of information gathering, the underlying assumption seems to be that the actual information will come through attorneys filling out some type of written forms.⁹⁰ Again, the Concept Paper treats the method of gathering the information as an open question.⁹¹ Further, the Concept Paper recognizes that, depending upon both the purposes and the criteria, information may be desired which cannot be obtained from such polls of attorneys.⁹²

Additional Matters. Finally, the Concept Papers calls for consideration of certain matters not addressed by *Judicial Performance Polls* including the following:

- (1) The development of specific standards for use in evaluating judges;
- (2) The development of a "kit" of materials which could be readily adopted by local associations for use in evaluation;
- (3) Consideration of whether the raw data should be weighed or refined in any way prior to release; and
- (4) Consideration of the frequency with which evaluations should be conducted.

This comparison should make it apparent that the focus of the inquiry of Maddi's work is much different than that of the proposed Special Committee on the Evaluation of Judicial Performance. Because both the object and the scope of inquiry are so different, the work of the proposed Special Committee would not be duplicative; to the contrary, even building upon Ms. Maddi's significant contributions, the task set forth for the Special Committee would be a large one.

IV. WHAT CRITERIA SHOULD BE UTILIZED?

The major substantive issue facing anyone working on a program to evaluate judicial performance is what standards, criteria, or "indicators" should be utilized. As former California Chief Justice Roger Traynor has recognized, "the difficulties of evaluating job performance at a workbench where each job is unique" are particularly difficult ones.⁹³ In addition, these problems are ones which may vary according to the purpose of the evaluation program.⁹⁴ Though it is certainly open to debate and consideration,

90. See generally D. MADDI, *supra* note 51, at 15-22.

91. CONCEPT PAPER, *supra* note 47, at 7-8.

92. *Id.* at 6.

93. Traynor, *Transatlantic Reflection on Leeways and Limits of Appellate Courts*, 1980 UTAH L. REV. 255.

94. CONCEPT PAPER, *supra* note 47, at 13.

the Evaluation of Judicial Performance Committee of the A.B.A. Criminal Justice Section has proposed that any A.B.A. inquiry into this area should specifically eschew any attempt to direct evaluation of performance at retention, selection, or discipline objectives.⁹⁵ Rather, the Evaluation Committee has proffered as the overriding purpose of such a program an educational objective: the improvement of sitting judges already upon the bench.⁹⁶

If this goal is accepted as a proper one, consideration should be given to two overriding issues with respect to what criteria should be used to evaluate judges' performance. First, attention should be focused on the standards which a "good" judge should strive to meet. Second, once those standards are identified, a judge's progress in striving to meet those standards must be measured or evaluated.

The first issue is one on which there is a wealth of data and scholarly writings, based both in theory and practice. In spite of the difference between a systematic form of judicial evaluation and bar selection and retention polls, the latter focuses upon the qualities that make up a "good" judge. They therefore offer a wealth of information upon which consideration of standards can be based.

Using a trial setting as a point of focus,⁹⁷ the Evaluation Committee has suggested that among the general criteria to be considered, the following might be included: (1) legal ability and knowledge; (2) diligence; (3) interpersonal traits; (4) judicial temperament and integrity; (5) conduct outside the courtroom; (6) comprehension of the applicable law in a given case; (7) willingness to consider novel theories and ability to understand such ideas; (8) consideration of briefs and arguments in an area of law which may be previously undecided or unfamiliar to the judge; (9) attitudes toward counsel and litigants; (10) industry; (11) judicial temperament—patience, courtesy, sense of humor, courage and dignity; (12) appearance of fairness and impartiality; and (13) actual fairness and impartiality.⁹⁸

95. *Id.* at 9.

96. *Id.* at 10.

97. The Evaluation Committee was particularly conscious of the fact that because of differing functions, different standards might well be necessary for appellate judges, administrative law judges, and others. CONCEPT PAPER, *supra* note 47, at 12.

98. *Id.* at 13.

Though these criteria were not intended to be exclusive and were offered only for discussion purposes, they seem to be consistent with the standards utilized in the selection and retention process in a number of other contexts, including the United States Circuit Judge Nominating Commission,⁹⁹ the Chicago Bar Association Guidelines for Judicial Selection which were adopted by the Lawyers Conference, Judicial Quality and Performance Committee of the A.B.A. Judicial Administration Division,¹⁰⁰ the

99. Both of President Carter's executive orders included the following as criteria for selection of a circuit judge:

Sec. 4. Standards for Selection of Proposed Nominees.

(a) Before transmitting to the President the names of the persons it deems best qualified to fill an existing vacancy or vacancies, a panel shall have determined:

(1) That those persons are members in good standing of at least one state bar, or the District of Columbia bar, and members in good standing of any other bars of which they may be members;

(2) That they possess, and have reputations for, integrity and good character;

(3) That they are of sound health;

(4) That they possess, and have demonstrated, outstanding legal ability and commitment to equal justice under law;

(5) That their demeanor, character, and personality indicate that they would exhibit judicial temperament if appointed to the position of United States Circuit Judge.

(b) In selecting persons whose names will be transmitted to the President, a panel shall consider whether the training, experience, or expertise of certain of the well-qualified individuals would help to meet a perceived need of the court of appeals on which the vacancy or vacancies exist.

Exec. Order No. 11,972, 42 Fed. Reg. 9,659 (1977) and Exec. Order No. 12,059, 43 Fed. Reg. 20,949 (1978) *reprinted in* L. BERKSON & S. CARBON, *supra* note 27 at 211 and 214. These standards were amplified by the Department of Justice, U.S. DEP'T. OF JUSTICE, SUPPLEMENTAL INSTRUCTIONS (1978) *reprinted in* L. BERKSON & S. CARBON, *supra* note 27, at 217 and 221 respectively.

100. Memorandum from James P. Economos, Chairman, Lawyers Conference Judicial Quality and Performance Committee, of the A.B.A. Judicial Administration Division, to the members of the Committee (undated). The Criteria set forth are: (1) integrity; (2) legal knowledge and ability; (3) judicial temperament; (4) diligence; (5) punctuality; (6) health; (7) age; (8) professional experience; (9) litigation experience; (10) past professional conduct; (11) financial responsibility; (12) political activity; (13) character; (14) patience; (15) common sense; (16) tact; (17) social consciousness; and (18) association and public service.

One of the particularly admirable features of the Chicago Bar Association Guidelines is the thoughtful attempt to define these terms. For illustrative purposes, consider one of the five paragraphs setting forth the meaning of legal knowledge and ability:

It is difficult to separate the concepts of legal knowledge and legal ability. Legal knowledge, in its simplest form, may be defined as familiarity with established legal concepts and procedural rules. Legal ability may similarly be defined as the intellectual capacity to interpret and apply established legal concepts to the facts and circumstances presented. Legal ability would also seem to involve skill in communicating, orally and in writing, and thought processes leading to a legal conclusion.

CHICAGO BAR ASSOCIATION, GUIDELINES FOR JUDICIAL SELECTION 3. These definitions have at least two beneficial aspects. It forces one to think with more specificity about the rating to be given a judge, and it should produce a greater uniformity in the use of terms and standards of evaluation.

model bar poll distributed by the A.B.A. Young Lawyer's Division,¹⁰¹ polls used by various local bar associations¹⁰² and the cri-

101. The Judicial Selection Committee of the A.B.A. Young Lawyers Committee developed a model bar poll which included the following criteria:

A. JUDICIAL CONDUCT

1. appearance of impartiality and freedom from:
 - a. racial bias
 - b. ethnic bias
 - c. sexual bias
 - d. political bias
 - e. economic bias
2. attentiveness to proceedings
3. courteousness to litigants and counsel
4. use of the consistent standards in accommodating time demands of lawyers
5. appearance of fairness to persons who are unrepresented by counsel
6. does health or age interfere with discharge of duties? (Yes[Y] or No [N])
7. does the judge refrain from ex parte communications in contested matters? (Y or N)

B. LEGAL ABILITY

1. quality and clarity of decisions
2. understanding of issues in complex cases
3. knowledge of and correct application of substantive law
4. knowledge of and correct application of the procedural law
5. keeps abreast of legal developments

C. COURTROOM MANAGEMENT

1. directs proceedings efficiently and with appropriate firmness
2. standard for thoroughness within the parameters of time
3. convenes court punctually
4. availability in chambers when court is not in session
5. reaches decision in reasonable amount of time
6. keeps docket as current as possible

HOW WOULD YOU RATE THE JUDGE'S OVERALL PERFORMANCE?

Attachment to Letter from David C. Weiner, *supra* note 56. Those given the poll are asked to indicate on a grid whether the judge should be rated excellent, good, adequate, needs improvement, or poor.

102. For example, the Arkansas Bar Association criteria for trial judges consists of twenty-three items: (1) character and integrity; (2) knowledge and application of substantive law; (3) knowledge and application of rules of evidence and of procedure; (4) comprehension of issues in highly complex cases; (5) efficient and conscientious worker; (6) judicial temperament; (7) compassion; (8) awareness of recent legal developments; (9) pre-trial and docket management-control; (10) participation in settlement efforts to a desirable degree; (11) punctuality in open court and making decisions; (12) unaffected by the identity of lawyers involved; (13) impartiality and fairness toward litigants and lawyers; (14) patience and courtesy to litigants, witnesses, jurors and lawyers; (15) trial management-control; (16) absence of bias or prejudice in criminal cases; (18) absence of bias or prejudice in domestic relations cases; (19) absence of bias or prejudice generally (*e.g.*, religion, race, sex, economic); (20) absence of political considerations and decisions; (21) not influenced by improper approaches; (22) promptness in making and giving decisions, and (23) quality of written opinions. ARKANSAS BAR ASSOCIATIONS, JUDICIAL POLL, 4-6 (1980). The Los Angeles County Bar Association Special Committee

teria utilized by the news media.¹⁰³

Indeed, an analysis of the fourteen surveys and polls reproduced in Maddi's *Judicial Performance Polls*¹⁰⁴ suggests that historically the attributes of a good judge can be grouped in thirty-three different categories¹⁰⁵ closely related to those outlined by the Evaluation Committee. These categories are: the judge's age and health;¹⁰⁶ judicial temperament;¹⁰⁷ courage;¹⁰⁸ character and

on Judicial Evaluation utilized ten criteria in its 1976 evaluation of candidates for Municipal Judges: (1) integrity and character; (2) judgment and intellectual capacity; (3) experience, including but not limited to trial experience; (4) industry and diligence; (5) judicial temperament, including whether the candidate would be courteous and considerate of counsel, parties, witnesses and jurors and whether the candidate was even-tempered; (6) professional ability and knowledge of the law; (7) health; (8) general reputation in the community; (9) civic and community activities, and (10) any other relevant matters of concern. L.A. COUNTY ASSOCIATION, SPECIAL COMMITTEE ON JUDICIAL EVALUATION, EVALUATION OF MUNICIPAL JUDGE CANDIDATES (1976).

103. *E.g.*, Atkinson & Stiteler, *The Good, the Bad and the Downright Dangerous*, D (Dallas) MAGAZINE, Aug. 1979, at 64. This article indicates that the authors gave consideration to the following:

1. Statistical performance in the disposition of cases, which was said to be related to hard work.
2. The Dallas Bar poll which rated judges in ten categories from punctuality and courtroom decorum to knowledge of the law.
3. Interviews with 30 attorneys selected by the magazine, seeking information on:
 - a. the judge's knowledge of the law;
 - b. judicial temperament, including bias, courtesy, individualized decision-making; and
 - c. willingness to make hard decisions.

Id. at 64-65.

104. D. MADDI, *supra* note 51.

105. Obviously through this section, the categories may overlap and reasonable people can argue over the classification system. Acknowledging this fact, it is nevertheless submitted that the classification system provides a useful tool in assessing the extent to which there may be a consensus among the various polling organizations as to the qualities that make for a good judge.

106. The State Bar of Arizona questionnaire asked: "Is the judge's age and health such that the judge can effectively discharge the duties of judicial office?" The rating was yes or no. D. MADDI, *supra* note 51, at 34. The Barristers of the Beverly Hills Bar Association questionnaire included the following: "The judge's age or health does not interfere with the adequate discharge of his duties." *Id.* at 42. The rating was from strongly agree to strongly disagree. The Dallas Bar Association questionnaire asked: "Is his age such that he can effectively discharge the duties of his office?" The rating was yes or no. *Id.* at 57.

107. The State Bar of Arizona questionnaire included the category of: "[j]udicial temperament and demeanor." The rating was from excellent to very poor. *Id.* at 34.

The Barristers of the Beverly Hills Bar Association questionnaire had a similar category: "The judge is temperamentally suited to his position." The rating was from strongly agree to strongly disagree. *Id.* at 43.

The Chicago Council of Lawyers questionnaire included the following:

Judicial Temperament

He/She listens patiently to substantial arguments from all sides.

He/She is courteous towards lawyers and litigants.

He/She conducts court proceedings with dignity.

reputation;¹⁰⁹ courteousness;¹¹⁰ lack of arrogance;¹¹¹ patience, un-

He/She conducts court proceedings with appropriate firmness.
He/She gives due consideration to the convenience of lawyers and litigants in scheduling proceedings.
He/She refrains from prejudging the outcome of a case during pretrial or early trial proceedings.
He/She refrains from coercing settlements in compliance with his/her views.

Id. at 50. The rating was from strongly agree to strongly disagree. The Dallas Bar Association questionnaire asked: "Does he possess and demonstrate a proper judicial temperament and demeanor?" The rating was yes or no. *Id.* at 57.

The Houston Bar Association questionnaire asked: "Does he exhibit a judicial temperament in the courtroom?" The rating was yes or no. *Id.* at 61.

The Maryland State Bar Association questionnaire included several areas. "Temperament: Listens patiently and courteously to lawyers and litigants on both sides. Conducts court proceedings with appropriate firmness. Indicates an opinion on the outcome of the case before all evidence is submitted. Coerces settlements in compliance with own views." *Id.* at 66. The rating was from frequently to not at all.

The Bar Association of San Francisco questionnaire had a number of categories. "Judicial Temperament and Demeanor: Demeanor with which court proceedings are conducted. Tolerance and self-control. Courtesy to counsel, witnesses and litigants. Attentiveness to, and patience with, questioning and arguments of counsel and testimony of witnesses. Restraint from usurping the role of competent counsel in questioning witnesses. Judicial temperament and demeanor in general." *Id.* at 80. The rating was from excellent to unsatisfactory.

The Santa Clara County Bar Association questionnaire included "Temperament." *Id.* at 85. The rating was from excellent to poor.

The Alaska Judicial Council questionnaire included "Dignity of demeanor on the bench." The rating was from poor to excellent. The questionnaire also included another category. "GENERAL CHARACTERISTICS: Dignity of demeanor on the bench. Conducts self in a manner free from impropriety or the appearance of impropriety. Proper accessibility to counsel for special proceedings, motions for extraordinary relief, bail review, and the like." *Id.* at 102. The rating was from poor to excellent.

The State of Utah Judicial Qualification Commission questionnaire asked: "Does he possess a proper judicial temperament and demeanor?" *Id.* at 107. The rating was yes or no.

108. The Bar Association of San Francisco questionnaire included "Courage." *Id.* at 72. The rating was from excellent to poor.

109. The Houston Bar Association questionnaire asked: "Is he a man of good character and reputation?" *Id.* at 61. The rating was yes or no.

110. The State Bar of Arizona questionnaire included "Courteousness to litigants, witnesses, jurors, and lawyers." *Id.* at 35.

The Barristers of the Beverly Hills Bar Association questionnaire included "The judge is courteous to witnesses." *Id.* The rating was from strongly agree to strongly disagree.

The Dallas Bar Association questionnaire asked: "Is he courteous to counsel?" "Is he courteous toward counsel, litigants and witnesses?" *Id.* at 57. The rating was yes or no.

The Bar Association of San Francisco questionnaire asked about "Courtesy." *Id.* at 72. The rating was from excellent to poor.

111. The Santa Clara County Bar Association questionnaire asked about: "Humility and Compassion." *Id.* at 85. The rating was from excellent to poor.

derstanding, and compassion;¹¹² sense of humor;¹¹³ punctuality and promptness;¹¹⁴ conscientiousness;¹¹⁵ diligence and industry;¹¹⁶ fairness;¹¹⁷ impartiality;¹¹⁸ integrity and lack of bias;¹¹⁹

The Alaska Judicial Council questionnaire included: "Freedom from arrogance." *Id.* at 97. The rating was from poor to excellent.

112. The Bar Association of San Francisco questionnaire included: "Humanity and Compassion." *Id.* at 72. The rating was from excellent to poor.

The Santa Clara County Bar Association questionnaire included: "Humility and Compassion." *Id.* at 85. The rating was from excellent to poor.

The Alaska Judicial Council questionnaire had several categories. "Human understanding and compassion." "Patience, tolerance, self-control." "Talent and ability for children's cases." *Id.* at 97, 102. The ratings were all from poor to excellent.

113. The San Bernadino County Bar Association questionnaire included "Sense of Humor." *Id.* at 72. The rating was from excellent to poor.

114. The State Bar of Arizona questionnaire included: "Punctuality" and "Promptness in making rulings and in rendering decisions." *Id.* at 35. The rating was from excellent to very poor.

The Dallas Bar Association questionnaire asked "Is he punctual in opening Court and keeping appointments such as pre-trials, motion hearings and consultations with counsel?" "Is he prompt in making rulings and rendering documents?" *Id.* at 57. The ratings were yes or no.

The Houston Bar Association questionnaire asked: "Does he render prompt decisions?" *Id.* at 61. The rating was yes or no.

The Alaska Judicial Council questionnaire had two areas. "Punctuality in opening court and keeping appointments." "Reasonable promptness in making rulings and rendering decisions." *Id.* at 102. The rating was from poor to excellent.

115. The Alaska Judicial Council questionnaire included: "Conscientiousness in finding facts and/or interpreting the law without regard to possible public criticism." *Id.* The rating was from poor to excellent.

116. The Chicago Council of Lawyers questionnaire included: "Diligence: He/She convenes court punctually. If necessary he/she is willing to devote time beyond the working day to court business. His/Her hearings and pretrial conferences reflect adequate research and preparation." *Id.* at 50. The rating was from strongly agree to strongly disagree.

The Maryland State Bar Association questionnaire included the following:

Diligence:

Convenes court punctually.

Not absent from Court House during normal working hours, either part or all of a day.

Usually available in chambers for court business when not on bench during normal working hours.

When necessary, works beyond normal working hours, either on the bench, in chambers, or at home.

Rules promptly on pretrial motions.

Rules promptly on motions or objections made during trial.

Insures steady progress of a case during trial.

Decides cases promptly.

Id. at 66. The rating was from very good to very poor.

The Bar Association of San Francisco questionnaire included: "Industry and Promptness: Punctuality in opening court and keeping appointments. Promptness in making rulings and decisions *during* trial. Promptness in rendering decisions *after* trial. Willingness to work hard not adjourning or recessing court for personal convenience. Industry and promptness (in general)." *Id.* at 80. (emphasis in original). Rating was from excellent to unsatisfactory.

The Alaska Judicial Council questionnaire included: "Willingness to work diligently." *Id.* at 102. Rating was from poor to excellent.

117. The State Bar of Arizona questionnaire included: "Fairness toward all litigants." *Id.* at 34. Rating was from excellent to very poor.

The Dallas Bar Association questionnaire asked: "Does he act fairly toward all litigants and lawyers?" *Id.* at 57. Rating was yes or no.

The Houston Bar Association questionnaire asked: "Is he partial to any individual or group of lawyers or litigants?" *Id.* at 61. Rating was yes or no.

The Alaska Judicial Council questionnaire had two categories. "Sense of basic fairness and justice." "Apparent fairness and equality of treatment to all parties." *Id.* at 102, 104. Rating was from poor to excellent.

118. The Barristers of the Beverly Hills Bar Association questionnaire included the following:

The judge's conduct is free of bias based on sex. [Rating was from strongly agree to strongly disagree.]

The judge's conduct is free of bias based on race. [Rating was from strongly agree to strongly disagree.]

The judge's predisposition, if any, in class action suits is: [Rating: From strongly against to strongly in favor.]

The judge's predisposition, if any, in criminal cases, reflected by his actions and demeanor toward defendants, is: [rating was from strongly to innocence to strongly to guilt.]

The judge's political or other personal beliefs have no effect on his decisions. [Rating was from strongly agree to strongly disagree.]

The judge shows no favoritism toward individual attorneys. [Rating was from strongly agree to strongly disagree.]

Id. at 41-43.

The Dallas Bar Association questionnaire asked: "Is he impartial in rendering decisions?" Rating was yes or no. *Id.* at 57.

The Houston Bar Association questionnaire asked: "Does he allow bias or prejudice to influence him?" Rating was yes or no. *Id.* at 61.

The Maryland State Bar Association questioned bias as follows:

In criminal cases, displays favoritism toward either the prosecution or defendant.

In civil tort cases, displays favoritism toward either plaintiff or defendant.

Displays religious, racial or ethnic bias.

Displays sex bias against women.

Displays sex bias against men.

Rulings are affected by the identity of the lawyers involved.

Discusses pending matters with counsel of record ex parte without notifying all parties properly.

Can be influenced by improper approaches (personal, political, or financial).

Id. at 66. Rating was from frequently to not at all.

The Bar Association of San Francisco questionnaire used "Impartiality." *Id.* at 80. Rating was from excellent to poor.

The Bar Association of San Francisco questioned bias by the following:

Impartiality:

Conscientiousness in ruling as he/she interprets the law, without regard to possible appellate reversal or public criticism.

Avoidance of sexual, racial, or ethnic bias.

Restraint from favoritism toward the *prosecution* in *criminal* cases.

Restraint from favoritism toward the *plaintiff* in *civil torts* cases.

Restraint from favoritism toward the *defense* in *civil torts* cases.

Restraint from prejudging the outcome of the case.

Impartiality (in general).

Id. (emphasis in original). Rating was from excellent to unsatisfactory.

The Santa Clara County Bar Association questionnaire included "UNBIASED (In the Cultural and Racial Sense)." *Id.* at 85. Rating was from excellent to poor.

legal ability;¹²⁰ knowledge of law;¹²¹ preparation for case;¹²² atten-

The Alaska Judicial Council questionnaire had several criteria. "Equal treatment of all parties regardless of race, ethnic background, sex, social or economic status, or the like. Restraint from favoritism toward *either* prosecution or defense in criminal cases. Restraint from favoritism toward *either* plaintiff or defendant in civil cases. Restraint from prejudging outcome of the case." *Id.* at 102 (emphasis in original). Rating was from poor to excellent.

The State of Utah Judicial Qualification Commission questionnaire asked: "Do you approve, in general, of the manner in which he has conducted the affairs of his Court?" *Id.* at 107. Rating was yes or no.

119. The State Bar of Arizona questionnaire asked: "Does this judge have sufficient integrity to carry out the duties of judicial office?" *Id.* at 35. Rating was yes or no.

The Barristers of the Beverly Hills Bar Association questionnaire had several categories. "The judge's treatment of lawyers from large firms reflects:" Rating was from strong favoritism to strong disfavor. "The judge's treatment of large corporate defendants reflects:" Rating was from strong favoritism to strong disfavor. *Id.* at 42.

The Chicago Council of Lawyers questionnaire had the following:

Integrity:

His/Her rulings are uninfluenced by the identity of the lawyers and parties involved.

His/Her rulings in criminal cases are free from any predisposition to decide for either government or the defense.

His/Her appointments of trustees, receivers, masters, guardians, and other persons receiving fees for such appointments are made solely on the basis of merit.

He/She assures that fees and costs charged by trustees, receivers, masters, guardians, and similar appointees are fair and reasonable.

He/She refuses to discuss pending matters unless all parties have been properly notified.

His/Her decisions are free of influence from ex parte approaches of any nature.

Id. at 50. Rating was from strongly agree to strongly disagree.

The Houston Bar Association questionnaire asked: "Is he a man of honesty and integrity?" *Id.* at 61. Rating was yes or no.

The Santa Clara County Bar Association questionnaire had "JUDICIAL INTEGRITY (Commitment to the Law and the Judge's Code of Conduct)." *Id.* at 85. Rating was from excellent to poor.

The Alaska Judicial Council questionnaire had "Integrity: If you have rated any judge 'Poor' or 'Needs improvement' in this category, a brief explanatory statement in the space provided would be of assistance to the Alaska Judicial Council. This statement of reasons is entirely optional." *Id.* at 102. Rating was from poor to excellent.

The State of Utah Judicial Qualification Commission questionnaire asked: "Is he a man of good character and integrity?" Rating was yes or no.

120. The Chicago Council of Lawyers questionnaire used the following:

Legal Ability:

He/She understands the issues in highly complex cases.

He/She understands the issues in ordinary criminal cases.

He/She understands the issues in ordinary civil cases.

His/Her written rulings are clearly expressed.

His/Her oral rulings are clearly expressed.

He/She keeps abreast of legal developments.

Rating was from strongly agree to strongly disagree. In addition the questionnaire asked, "What is your legal specialization?" *Id.* at 50.

The Maryland State Bar Association questionnaire used several categories. "Legal Ability: Understands issues in highly complex cases. Competent in the usual civil cases. Competent in the usual criminal cases. Written rulings sound in sub-

tiveness to arguments;¹²³ consideration of written arguments;¹²⁴

stance. Oral rulings sound in substance. Keeps abreast of legal developments." *Id.* at 66. Rating was from very good to very poor.

The Bar Association of San Francisco questionnaire had similar areas. "Legal Skills: Knowledge of civil law: substantive, evidentiary, and procedural. Knowledge of criminal law: substantive, evidentiary, and procedural. Settlement skills. Familiarity with new legal developments. Legal skills (in general)." *Id.* at 80. [Rating: From excellent to unsatisfactory.]

The Alaska Judicial Council questionnaire had: "Legal reasoning ability and comprehension." *Id.* at 102. Rating was from poor to excellent.

121. The State Bar of Arizona questionnaire used several categories. "Knowledge of the law," "Knowledge and application of rules of evidence and substantive law," "Knowledge and application of rules of procedure." *Id.* at 34-35. Rating for each category was from excellent to very poor.

The Barristers of the Beverly Hills Bar Association questionnaire used several indicators. "The judge correctly instructs the jury in the law." "The judge knows and properly applies the rules of evidence." "The judge is well-versed in criminal law." "The judge is well-versed in civil law." *Id.* at 41-43. Rating for each category was from strongly agree to strongly disagree.

The Dallas Bar Association questionnaire asked: "Does he correctly apply the Rules of Procedure and the substantive law?" "Does he correctly apply the Rules of Procedure?" "Does he correctly apply the substantive law?" *Id.* at 57. Rating was yes or no.

The Houston Bar Association questionnaire asked: "Does he know and apply the law?" *Id.* at 61. Rating was yes or no.

The Bar Association of San Francisco questionnaire had "Legal knowledge." *Id.* at 72. Rating was from excellent to poor.

The Santa Clara County Bar Association questionnaire used several categories. "Intellect," "Knowledge (Procedure)," "Knowledge (Substantive Law)," and "Knowledge (Evidence)." *Id.* at 85. Rating for each was from excellent to poor.

The Alaska Judicial Council questionnaire included the following categories. "Knowledge of criminal substantive law, evidence and procedure." "Knowledge of civil substantive law, evidence and procedure." "Familiarity with available correctional programs, alternatives, and facilities." *Id.* at 97, 102. Rating for each was from poor to excellent.

The State of Utah Judicial Qualification Commission questionnaire asked: "Does he know and apply the Rules of Procedure?" "Does he have adequate knowledge of substantive law?" *Id.* at 107. Rating was yes or no.

122. The Barristers of the Beverly Hills Bar Association questionnaire asked whether "[t]he judge displays preparation for the case at hand." *Id.* at 39. Rating was from strongly agree to strongly disagree.

123. The State Bar of Arizona questionnaire included two areas. "Attentiveness to arguments of counsel." "Attentiveness to testimony of witnesses and arguments of counsel." *Id.* at 34-35. Rating was from excellent to very poor.

The Barristers of the Beverly Hills Bar Association questionnaire asked whether "[t]he judge is attentive to the proceedings." *Id.* at 39. Rating was from strongly agree to strongly disagree.

The Dallas Bar Association questionnaire asked: "Is he attentive to arguments of counsel?" "Is he attentive to testimony of witnesses and arguments of counsel?" *Id.* at 57. Rating was yes or no.

124. The State Bar of Arizona questionnaire asked about "[c]onsideration of briefs and authorities." *Id.* at 39. Rating was from excellent to very poor.

The Barristers of the Beverly Hills Bar Association questionnaire asked

quality of judge's decisions;¹²⁵ judge's rulings;¹²⁶ quality of written opinions;¹²⁷ willingness to learn;¹²⁸ courtroom management;¹²⁹ use of court time;¹³⁰ performance in conducting trials;¹³¹ firmness;¹³² decisiveness;¹³³ settlement skills;¹³⁴ judge's attitude toward negotiated pleas;¹³⁵ appropriateness of sentences in criminal cases;¹³⁶

whether "The judge gives due consideration to all motions and petitions." *Id.* at 39. Rating was from strongly agree to strongly disagree.

The Dallas Bar Association questionnaire asked "Does he carefully consider briefs and authorities submitted by counsel?"

"Does he carefully consider authority submitted by counsel?" *Id.* at 57. Rating was yes or no.

125. The Barristers of the Beverly Hills Bar Association questionnaire asked whether "[t]he judge's decisions are informed and aptly based on authority and policy." *Id.* at 40. Rating was from strongly agree to strongly disagree.

126. The Barristers of the Beverly Hills Bar Association questionnaire asked whether "[t]he judge's procedural rulings are prompt and proper." *Id.* Rating was from strongly agree to strongly disagree.

127. The State Bar of Arizona questionnaire had the category of "Quality of written opinions." *Id.* at 34. Rating was from excellent to poor.

The Dallas Bar Association questionnaire asked: "Are his written opinions of good quality?" *Id.* at 57. Rating was yes or no.

The Alaska Judicial Council questionnaire had several categories: "QUALITY OF WRITTEN OPINIONS: Legal Knowledge. Clarity and precision. Level of literacy and style. Restraint from favoritism toward *either* prosecution or defense in criminal cases. Restraint from favoritism toward *either* plaintiff or defendant in civil cases." *Id.* at 99. Rating was from poor to excellent.

128. The Bar Association of San Francisco questionnaire had the category of "Willingness to Learn." *Id.* at 72. Rating was from excellent to poor.

129. The State Bar of Arizona questionnaire had the area of "Courtroom discipline." *Id.* at 35. Rating was from excellent to very poor.

The Santa Clara County Bar Association questionnaire had a similar area: "Trial Management." *Id.* at 85. Rating was from excellent to poor.

The Alaska Judicial Council questionnaire had the category of "Ability to maintain proper control over courtroom." *Id.* at 102. Rating was from poor to excellent.

130. The Barristers of the Beverly Hills Bar Association questionnaire asked whether "[t]he judge makes effective use of court time." *Id.* at 40. Rating was from strongly agree to strongly disagree.

The Houston Bar Association questionnaire asked "Does he make appropriate use of his time in the courtroom?" *Id.* at 61. Rating was yes or no.

131. The Barristers of the Beverly Hills Bar Association questionnaire asked whether "[t]he judge allows counsel ample opportunity to present and develop arguments." *Id.* at 39. Rating was from strongly agree to strongly disagree.

The Dallas Bar Association questionnaire asked "Does he allow sufficient time to counsel to develop fully a case in trial?" *Id.* at 57. Rating was yes or no.

The Alaska Judicial Council questionnaire had several categories. "Performance in trial of cases." "Performance as civil motions judge (discovery, summary judgment, and the like)." "Performance as criminal motions judge (discovery, summary judgment, and the like)." *Id.* at 97. Rating was from poor to excellent.

132. The Bar Association of San Francisco questionnaire had the area of "Firmness." *Id.* at 72. Rating was from excellent to poor.

133. The Chicago Council of Lawyers questionnaire had several categories. "Decisiveness [:] He/she rules promptly on pretrial motions. He/she insures steady progress of a case prior to trial. He/she rules with appropriate decisiveness during trial." *Id.* at 50. Rating was from strongly agree to strongly disagree.

134. The Bar Association of San Francisco questionnaire included the area of "Settlement Skills." *Id.* at 72. Rating was from excellent to poor.

135. The Barristers of the Beverly Hills Bar Association questionnaire asked

quality of judge's instruction,¹³⁷ and overall quality as a judge.¹³⁸ This analysis corresponds with a 1977 study of the criteria utilized by twenty-five bar associations surveyed.¹³⁹

Thus, without intending to deny that reasonable people can disagree upon the general criteria to be utilized or to suggest that this is not a matter for serious inquiry and discussion, there is substantial agreement upon many of the standards by which judges could be evaluated. This being the case, it is entirely possible that a hard-working group of diverse individuals could come to a consensus upon the criteria to be utilized in evaluating judicial performance.

whether "[t]he judge sanctions and supports negotiated pleas." *Id.* at 40. Rating was from strongly agree to strongly disagree.

136. The Barristers of the Beverly Hills Bar Association questionnaire had two categories. "The judge's sentences for violent crimes are:" Rating was from very lenient to very severe. "The judge's sentences for white collar defendants are:" Rating was from very lenient to very severe. *Id.* at 43.

The Alaska Judicial Council questionnaire had the area of "Consistency in sentencing practices." *Id.* at 102. Rating was from poor to excellent.

137. The Barristers of the Beverly Hills Bar Association questionnaire asked whether "[t]he judge's instructions to the jury are delivered in a non-prejudicial manner." *Id.* at 40. Rating was from strongly agree to strongly disagree.

138. The Chicago Council of Lawyers questionnaire asked "Overall, is he/she worthy of advancement to a higher judicial office?" *Id.* at 50.

The Barristers of the Beverly Hills Bar Association questionnaire asked whether "[t]he judge is an outstanding member of the federal judiciary." *Id.* at 43. Rating was from strongly agree to strongly disagree.

The State Bar of Arizona questionnaire asked "Should this judge be retained in office?" *Id.* at 35.

The Chicago Council of Lawyers questionnaire asked "Overall, is he/she worthy of retention in his/her present post?" *Id.* at 50.

The Dallas Bar Association questionnaire asked: "Do you approve in general of the overall manner in which this Judge has conducted the affairs of his Court?" *Id.* at 57.

The Maryland State Bar Association questionnaire asked: "In light of your answers above, how do you evaluate the judge's performance?" *Id.* at 66.

The Bar Association of San Francisco questionnaire asked about "Your Overall Rating of Each". *Id.* at 72. Rating was from excellent to poor.

The Santa Clara County Bar Association questionnaire had the category of "Overall Rating." *Id.* at 85. Rating was from excellent to poor.

The Alaska Judicial Council questionnaire asked about "[o]verall judicial performance." *Id.* at 102. Rating was from excellent to poor.

The State of Utah Judicial Qualification Commission questionnaire asked whether approval could be given of the manner, in general, in which the judge had conducted the court's affairs. *Id.* at 107. Rating was yes, no, or no opinion.

139. J. GUTERMAN & E. MEIDINGER, *supra* note 58, at 38-41, grouped the criteria utilized as follows:

The more difficult question is how is the progress of a judge who is striving to meet those standards to be evaluated. In a study of the performance of prosecutors, defense attorneys, and judges in matters of case screening, case efficiency and delay, plea bargaining, and sentence variation, considerable attention was given to finding measures which could be used to judge per-

CRITERIA	N	PERCENT
<u>Technical Qualifications</u>		
Legal Ability	18	72.0
Legal Knowledge	7	28.0
Legal Experience	5	20.0
Quality of Opinions	5	20.0
Procedural Correctness	4	16.0
Substantive Correctness	2	8.0
Evidentiary Correctness.....	1	4.0
Intellect	1	4.0
<u>Work Capacity</u>		
Diligence/Industry	17	68.0
Punctuality/Promptness	13	52.0
Trial Management	6	24.0
Studiousness	6	24.0
Settlement Skills	4	16.0
Age	4	16.0
Administrative Skill.....	2	8.0
Efficiency	1	4.0
Physical/Mental Fitness.....	1	4.0
<u>Interactive Traits</u>		
Courtesy	16	64.0
Attentiveness	7	28.0
Proper Demeanor	4	16.0
Lacking Controversial Conduct	4	16.0
Patience	4	16.0
Considerateness	3	12.0
Respect for Lawyers	2	8.0
Sense of Humor	2	8.0
<u>Character Traits</u>		
Judicial Temperament	18	72.0
Integrity	13	52.0
Impartiality	12	48.0
Lack of Bias/Prejudice	10	40.0
Political/Economic Independence	5	20.0
Decisiveness/Firmness	4	16.0
Courage	3	12.0
Intellectual Honesty	3	12.0
General Character Fitness	2	8.0
Judgment/Perspective	2	8.0
Neutrality.....	2	8.0
Willingness to Learn	2	8.0
Additional Questions on General Qualifications for Office	17	68.0

formance.¹⁴⁰ As a general guideline, it was concluded that such "performance measures" should be relevant,¹⁴¹ specific,¹⁴² and clear.¹⁴³ Further, the more proximate,¹⁴⁴ directly linked,¹⁴⁵ and

140. S. WILDHORN, M. LAVIN & A. PASCAL, LAW ENFORCEMENT ASSISTANCE ADMINISTRATION, INDICATORS OF JUSTICE: MEASURING THE PERFORMANCE OF PROSECUTION, DEFENSE AND COURT AGENCIES INVOLVED IN FELONY PROCEEDINGS, A GUIDE TO PRACTITIONERS (1977). See also S. WILDHORN, M. LAVIN & A. PASCAL, LAW ENFORCEMENT ASSISTANCE ADMINISTRATION, INDICATORS OF JUSTICE: MEASURING THE PERFORMANCE OF PROSECUTION, DEFENSE AND COURT AGENCIES INVOLVED IN FELONY PROCEEDINGS, ANALYSIS AND DEMONSTRATION (1977) [hereinafter cited as ANALYSIS AND DEMONSTRATION].

141. ANALYSIS AND DEMONSTRATION, *supra* note 140, at 33-34.

A *relevant* performance measure is one with significant probative value concerning a matter that is in issue or that requires illumination. For example, the average elapsed time from arraignment to disposition is a measure relevant to issues of court resource use. The proportion of cases dismissed before trial is a measure that has little relevance to issues of sentence variation.

Id. at 33-34 (emphasis in original).

142. *Id.* at 34.

The fewer the aspects of the proceeding on which the informational content of a performance measure is focused, the more *specific* it is. For example, case rejection rate is a measure more specific to the screening process, on which it is singularly focused, than is a measure such as the proportion of cases in which the defendant pleads to charges different from the ones originally filed, which reflects both the charging and plea bargaining aspects.

Id. at 34 (emphasis in original).

143. *Id.*

Clarity connotes that the construction and usage of a performance measure is readily understood by the average practitioner. For example, median number of days between arrest and final disposition of all felony cases is a clear measure of delay. However, a sentence severity 'score' is a less clear performance measure of sentence severity than the elements that comprise it (e.g. two years of prison followed by five years on probation).

Id. (emphasis in original).

144. *Id.*

A performance measure may be more or less *proximate*, depending on how close the events it captures are to the matter of interest. For example, the proportion of complaints rejected by the prosecutor for reasons of insufficient evidence is an immediate output of screening and is therefore proximate to the matter of screening performance. By contrast, the rate at which trial dismissals occurs is a measure of events more remote from the screening process and is therefore less proximate to the prosecutor's screening performance.

Id. (emphasis in original).

145. *Id.* "A performance measure is more *directly linked* to an aspect of the criminal proceeding if its magnitude is more strongly correlated with that aspect. For example, the guilty plea rate would be more directly linked with judicial case processing efficiency than would be, say, a measure of sentencing severity." *Id.* (emphasis in original).

applicable¹⁴⁶ the measure is, the greater its reliability.

One solution to this problem of judging a jurist's improvement is to conclude that the traditional method of soliciting the opinions and impressions of attorneys as to the judge's performance in each of these areas is a satisfactory method of evaluation. While this approach would be subject to the infirmities set forth previously,¹⁴⁷ it would, especially if attempts were made to strengthen the objectivity of the questions and to improve the response rate, provide data that could be useful to the evaluation of judicial performance.

Another solution, which could be pursued either in conjunction with a survey of attorneys or independently, is to attempt to find more specific and objective criteria which would indicate the quality of a given judge's performance. For example, it might provide a more reliable guide to assessing a judge's integrity if the impressions of the bar were compared with any disciplinary action that might have been sought or obtained against that judge. In an attempt to isolate such specific criteria which could be used to assess judicial performance, the Evaluation Committee proffered the following examples for consideration for use at the trial level:

- (1) frequency of reversal on appeal;
- (2) number of cases handled over a period of time (three years, for example);
- (3) types of cases handled (nature of cases, jury trial, etc.);
- (4) time between submission of a case and decision;
- (5) number of cases settled before trial;
- (6) number of cases settled during trial;
- (7) hours of attendance at continuing education courses;
- (8) number of postponements of hearings, conferences, etc.,
- (9) sentencing data;
- (10) number of complaints filed with judicial disciplinary agencies;
- (11) frequency of complaints with pertinent rules (such as flings of Findings of Facts, etc.);
- (12) data concerning movement of the docket (as well as study of content of the daily docket);
- (13) disparity of sentencing as compared to other judges of the same court and in the same court system;
- (14) industry or the amount of time devoted to judicial duties and in furtherance of the administration of justice;

146. *Id.*

The *applicability* of a performance measure refers to its usefulness in the analytical task undertaken. For example, if one's purpose is to make inter-jurisdictional comparisons of judicial productivity, the number of *weighted* cases processed per available judge-year may be a more useful measure for this purpose than would be simply the *unweighted average* number of disposition per judge per year.

Id. (emphasis in original).

147. See notes 56-66 *supra* and accompanying text.

- (15) character, honesty, integrity and maturity;
- (16) participation and/or invitations to participate as a lecturer at law schools, legal seminars, etc.¹⁴⁸

There are several benefits in utilizing such objective criteria. First, objective criteria are less subject to bias or distortion than subjective criteria. Second, if the objective criteria and subjective criteria coincide, they verify the reliability of each. If they conflict, that would suggest further inquiry is necessary. Either way, the reliability of the results is enhanced. Third, as a by-product of the foregoing, the entire evaluation itself becomes more reliable and more acceptable to both judges and those participating in the process.

Nevertheless, it should be recognized that even the objective criteria can also be distorted. One example would be the use of statistics of a trial judge's rate of being reversed by an appellate court. This should be a statistic which is fairly easy to compile. On its face, one might tend to think that those judges who had been reversed more times were not as good as those who had a better record on appeal. This raw statistic, however, could obviously be misleading. There are any number of cases in which there seems to be no statute or case law directly on point when a given issue must be decided. This requires the court to make a

148. Qualifiable standards for appellate courts might include: (1) conciseness of written opinions; (2) length of time from hearing and/or assessment of case to circulation of written opinion; (3) frequency of dissenting opinions; (4) number of cases handled over a period of time; (5) types of cases handled; (6) time between submission of a case and a decision; (7) hours of attendance at continuing education courses; (8) number of complaints filed with the judicial disciplinary agencies; (9) frequency of compliance with pertinent rules. CONCEPT PAPER *supra* note 47, at 13.

Compare with the criteria adopted for appellate courts by the ARKANSAS BAR ASSOCIATION, JUDICIAL POLL, 4-6 (1980):

1. Character and integrity.
2. Good knowledge of substantive law.
3. Efficient and conscientious worker.
4. Impartiality and fairness toward lawyers.
5. Awareness of recent legal developments.
6. Understands issues in highly complex cases.
7. Quality of written opinions.
8. Comprehension of significance and implication of judicial precedents.
9. Judicial temperament for a judge.
10. Awareness of legal and factual issues.
11. Not influenced by improper approaches.
12. Unaffected by the identity of the lawyers involved.
13. Absence of political considerations in discussions.
14. Absence of prejudice against party requesting oral argument.
15. Patience and courtesy to all lawyers.

decision based upon public policy as enunciated by prior analogous cases or analogous statutes. The fact that a judge may make a choice between two equally reasonable alternatives and be reversed by a higher court simply because that court happens to have made a different choice based upon their personal predispositions, should not in any way adversely affect the assessment of the performance of the lower court judge.¹⁴⁹

Similarly, there may be other instances because of changes in public policy or circumstances, where it may appear that existing precedent has been so eroded that the demands of current jurisprudence require that the lower court adopt new law recognizing that there was a possibility of being reversed upon appeal.¹⁵⁰ A third example might be instances where the lower court judge believed he/she was following firmly established precedent only to discover that the higher court itself would later abandon that precedent and reverse the lower court.¹⁵¹

Thus, while the raw statistics might be of interest in any evaluation of judicial performance, they would most assuredly be insufficient to provide an accurate picture of the performance of that judge without further analysis.

Similarly, a common measure of judicial performance is often thought to be whether judges keep their dockets current. Yet these statistics, too, may be misleading. For example, if a judge is assigned to a major school desegregation case,¹⁵² an anti-trust case,¹⁵³ or a civil rights suit,¹⁵⁴ that case may consume so much of

149. Consider the question of whether an incestuous common law marriage between uncle and niece was void or only voidable. If it was void, the wife would be precluded from claiming the right to serve as administratrix. The trial court held it was void, only to be reversed by a divided court of appeals. *In re Estate of David R. Stiles*, No. 1131 (Ohio Ct. App., Fourth District, August 2, 1978). In spite of the excellent analysis by the judge who authored the majority opinion, the decision of the court of appeals was reversed by the Ohio Supreme Court. *In re Estate of David R. Stiles*, 59 Ohio St.2d 73, 391 N.E.2d 1026 (1979).

150. *E.g.*, *Krause v. Ohio*, 28 Ohio App. 2d 1, 274 N.E.2d 321 (1971), *rev'd*, 31 Ohio St. 2d 132, 285 N.E.2d 736 (1972).

151. *E.g.*, *City of Kenosha v. Bruno*, 411 U.S. 904 (1973); *Moor v. County of Alameda*, 411 U.S. 693 (1973); *Monroe v. Pape*, 365 U.S. 167 (1961). *But see* *Monnell v. Department of Social Services of New York*, 436 U.S. 658 (1978). For example, in *Monnell*, the trial judge could justifiably rely upon three U.S. Supreme Court cases in determining that a county was not a person within the meaning of 42 U.S.C. § 1983 (Supp. 1979): *Monroe v. Pape*, 365 U.S. 167 (1961); *City of Kenosha v. Bruno*, 411 U.S. 904 (1973); and *Moor v. County of Alameda*, 411 U.S. 693 (1973). Yet the Supreme Court in *Monnell* found that *Pape* and its progeny were wrongly decided and held that a county was a person under § 1983.

152. *E.g.*, *Reed v. Rhodes*, 455 F. Supp. 546 (N.D. Ohio 1978), *modified*, 607 F.2d 714 (6th Cir. 1979) (Cleveland school desegregation suit).

153. *E.g.*, *The U.S. v. IBM: The Endless Suit*, *NEWSWEEK*, March 10, 1980, at 75-76.

154. *E.g.*, *Pugh v. Locke*, 406 F. Supp. 318 (M.D. Ala. 1976), *aff'd in part sub nom.*; *Newman v. Alabama*, 559 F.2d 283 (5th Cir. 1977), *cert. denied*, 438 U.S. 915 (1978) (Alabama prison conditions case).

one judge's time as to be the equivalent of several hundred more routine cases disposed of by other judges. Further, in those systems in which judges have certain flexibility in accepting or receiving assignments, or in which administrative judges have flexibility in making assignments, the number of cases processed may reflect on the difficulty of the cases. In addition, certain judges may go out of their way to look for reasons to disqualify themselves in particularly difficult cases. In jurisdictions where administrative judges have the power to assign particular cases to a particular judge, they may give the more difficult cases to those judges who may be considered to be more able or to have more expertise in that area. Also, some judges could use unfair settlement techniques in order to force cases which ought to go to trial into settlements in which neither of the parties are satisfied.¹⁵⁵

All this suggests that using raw statistics on case processing is not enough. Indeed, if mere statistics were used, then this might well provide incentive for judges to engage in some of the practices set forth above which would result in deterioration rather than improvement of judicial performance.

These illustrations are presented to indicate that the need to proceed with caution in using objective indicators is fully recognized. Whether the subject criteria, objective criteria, or some combination of the two is to be used in judicial evaluation presents a weighty issue which demands the serious attention of all interested parties. Hopefully the foregoing outline of some of the questions involved highlights the need for a thoughtful and sensitive inquiry into the standards and performance measures for evaluating judicial performance.

V. ISSUES IN METHODOLOGY¹⁵⁶

As the controversy over the reliability of bar polls should suggest, the methodology to be used in evaluating judicial performance raises questions over which reasonable people have honest

155. On the proper use of settlement in the federal courts, *see* FEDERAL JUDICIAL CENTER, *THE ROLE OF THE JUDGE IN THE SETTLEMENT PROCESS* (1977) which contains transcripts of remarks by Federal District Judges Will, Merhige and Rubin.

156. This section is taken, in large part, from the report of the Methodology Subcommittee of the A.B.A. Criminal Justice Section Evaluation of Judicial Performance Committee, of which the author was chairperson. This report was incorporated into the Concept Paper of the full committee. *CONCEPT PAPER*, *supra* note 47, at 3-9.

disagreement. The lack of consensus in this area is aggravated by the difficulties with existing systems of bar polling. Further, some of the methodological questions are so complex that any comprehensive and serious attempt at judicial performance evaluation will require the expertise of social scientists skilled in evaluation and survey techniques. It is, therefore, in this area that the proposed Special Committee may find its greatest challenge.

The objective here is not to recommend solutions to the differences of opinion that may exist. To the contrary, the purpose is to examine some of these issues in more detail; to raise additional questions; and to explore some—but obviously not all—of the options that could be given consideration in attempting to resolve them. In the broadest terms, these issues include:

- (1) Who should be given oversight responsibility for conducting such evaluation?
- (2) Who should actually conduct and administer such evaluations?
- (3) From what sources should information be gathered in the evaluation process?
- (4) How should the information be collected?
- (5) How frequently should such evaluations be conducted?
- (6) Should the raw data gathered be refined or “weighted” in any way?

A. Who Should be Given Oversight Responsibility for Conducting Such an Evaluation?

One of the critical questions involved in any discussion of evaluating judicial performance is who is going to run the program. That question may be subdivided into two separate questions. Who is going to be the sponsoring body, with overall responsibility for the program? Who is actually going to administer and implement the program? While it is possible that the same person or entity could function in both capacities,¹⁵⁷ it is also quite likely that in many circumstances each function would be performed by separate persons or entities.¹⁵⁸ For this reason, these questions will be addressed separately, recognizing that in some instances there will be an overlap of functions between those who have oversight responsibility and those who implement the program.

In attempting to consider all reasonable possibilities for sponsors of such an evaluation, reference was made to those groups which have sponsored selection or retention evaluations in the

157. It is not difficult to imagine a local evaluation system, in a court of less than ten judges, where a court administrator would act on behalf of the entire court in both capacities.

158. For example, in a state-wide program where supervisory authority was vested in a committee established by the supreme court, *e.g.*, the New Jersey plan, the actual implementation of the evaluation program must necessarily fall on other shoulders.

past. An attempt, however, was also made to consider others who may claim an interest in this process. Thus, the initial list of possibilities includes courts or some court-established institution;¹⁵⁹ bar associations, at a national, state, or local level; news media;¹⁶⁰ citizens groups; non-judicial branches of government; or some combination of the above. Obviously the selection of a sponsoring group for judicial evaluation is a sensitive question which must take into consideration who has the credibility and "clout" to secure the cooperation of the judges to be evaluated and who has financial and personnel resources to establish and maintain a quality program of judicial evaluation.

B. Who Should Actually Conduct and Administer Such Evaluations?

Once a decision is made as to who should have general oversight responsibility for judicial evaluation, the next question of methodology is who should be charged with the implementation of such a program. The options here raise a number of serious questions.

The initial issue to be confronted is the scope and extent to which such a program is to be implemented. Obviously, if an intention is to set up a mechanism for judicial evaluation for all judges across the entire country, the existence of a national entity to implement such a program would necessarily involve large bureaucratic problems. On the other hand, if the implementation is to be left to smaller subdivisions on a regional, state, or local basis, then there is an increased risk for failure, particularly where a

159. *E.g.*, Rubenstein, *Alaska's Judicial Evaluation Program: A Poll The Voters Rejected*, 60 JUDICATURE 478 (1977); D. MADDI, *supra* note 51, at 91-104 and 105-07 (Utah Judicial Qualification and Removal Commission); Handler, *supra* note 39, (New Jersey Supreme Court's Committee on Judicial Evaluation and Performance).

160. The Great Falls Tribune, of Montana, conducted its own poll of attorneys for use in evaluating judges in 1976. J. GUTERMAN & E. MEIDINGER, *supra* note 58, at 32, 58-78. Other newspapers embarking on independent polls have included those in Minneapolis, Minnesota; Portland, Oregon; and Seattle, Washington. D. MADDI, *supra* note 51, at 3, 3 n.4. In 1975, the Maryland State Bar Association and The Baltimore Sun jointly sponsored a poll of attorneys to evaluate 21 sitting judges in Baltimore. *See* PHILIP, *supra* note 58, at 36 and J. GUTERMAN & E. MEIDINGER, *supra* note 58, at 39. More recently polls have been abandoned for interviews and a journalistic approach in media evaluations of the judiciary. *E.g.*, Pike & Crosby, *Judging the Judges* (pts. 1-7). The Washinton Star, Jan. 9, 1978 to Jan. 14, 1978; Atkinson and Stiteler, *supra* note 103, at 64; R. WOODWARD AND S. ARMSTRONG, *THE BRETHREN* (1979).

substantial financial commitment is involved. An intermediate step perhaps would be to develop a system which could be implemented on the local level but would be developed nationally in a pre-packaged kit.¹⁶¹

Regardless of the scale upon which such evaluation will be attempted, the implementation of such a program raises numerous complex questions. Who will actually physically devise such an evaluation? Is expert assistance from those familiar with statistical and survey techniques necessary,¹⁶² and if so, to what degree? In the context of bar polls, this has been the most crucial and difficult task facing those administering the evaluation.¹⁶³ Indeed, in New Jersey, the Handler Committee concluded that the effective evaluation of judicial performance would entail a high order of skill, expertise and sophistication. It was also perceived that within the state there was no ready pool of personnel with suitable backgrounds or training to launch such a program.¹⁶⁴

If experts are utilized, to what degree do non-experts, such as judges and lawyers, participate in devising the techniques by

161. This is the approach the Florida Bar Association took with respect to polls. In 1974, it published its *Model Judicial Poll Handbook* for use by local bar associations. D. MADDI, *supra* note 51, at 86.

162. Even in bar polls, many associations have sought expert advice. Examples include:

POLLING GROUP	EXPERTS UTILIZED
1. Alaska Judicial Council	1. Center for Political Studies of the Institute for Social Research of the University of Michigan at Ann Arbor.
2. State Bar of Arizona	2. Survey Research Laboratory of the Department of Sociology of Arizona State University.
3. Barristers of Beverly Hills Bar Association	3. Ph.D. candidates at Claremont College.
4. Chicago Council of Lawyers	4. Political Science Laboratory of the University of Wisconsin—Milwaukee and the Department of Sociology of Northwestern University.
5. Bar Association of San Francisco	5. Professional Survey Team who were members of the Department of Sociology of San Francisco State University.
6. Maryland State Bar Association <i>The Baltimore Sun</i>	6. Sidney Hollander and Associates (an opinion research firm).

Rubenstein, *supra* note 159, at 481; D. MADDI, *supra* note 51, at 28, 36, 46, 62, 73.

163. J. GUTERMAN & E. MEIDINGER, *supra* note 58, at 35.

164. Handler, *supra* note 30, at 3.

which such an evaluation would take place?¹⁶⁵ If methodological disagreements arise between the social science experts and the lawyer or judge non-experts, how are these to be resolved?

A related question is whether laypersons, with no expertise in either law or the social sciences, should participate in the formation and implementation of a judicial performance evaluation program. Other than the Alaska experience, where members of the Alaska Peace Officers Association helped develop polls that were later circulated to determine whether incumbent judges should be retained in office,¹⁶⁶ there are no similar experiences to draw upon. The closest analogy may be to situations in which laypersons participate in the actual selection of judges, either directly or through an endorsement committee.

An example of direct participation by laypersons in the selection of judges is the United States Circuit Judge Nominating Commission. When this Commission was created in 1977, it was explicitly provided that "[e]ach panel shall include . . . approximately equal numbers of lawyers and nonlawyers."¹⁶⁷ A recent study analyzed the performance of that Commission and its panels. The existence of the panels spanned two periods of time. The first period covered the time prior to the passage of the Omnibus Judgeship Bill which created 152 new judgeships, and a second period¹⁶⁸ covered the time afterward.¹⁶⁹ Of the thirteen panels created, two panels in the first period and one in the second had a majority of non-lawyers.¹⁷⁰ Overall, forty-two percent of the members of the panels in the first round and thirty-eight in the second round were non-lawyers.¹⁷¹ Though the performance of the lay individuals on the Commission was not without criticism,¹⁷² most of the members of the Commission counted the

165. In the Alaska program, the Judicial Council invited delegates from the bar association and the Alaska Peace Officers Association to meet with them and participate in the design of the questionnaires to be used. Rubinstein, *supra* note 159, at 480.

166. *Id.*

167. Exec. Order No. 11,972, 42 Fed. Reg. 9,659 (1977) at § 2(c) (providing criteria for selection of circuit judges). *But cf.* Exec. Order No. 12,059, 43 Fed. Reg. 20,949 (1978) makes no specific provision for the appointment of laypersons, but § 2(c) provides that: "Each panel shall include . . . at least one lawyer from each State within a panel's area of responsibility."

168. L. BERKSON & S. CARBON, *supra* note 27, at 6.

169. *Id.*

170. *Id.* at 51.

171. *Id.*

172. The criticism generally focused upon the lack of knowledge on the part of

"good mix" of individuals as a positive factor and thought very highly of their fellow panelists.¹⁷³

In spite of criticism for partisanship,¹⁷⁴ an independent study funded by the American Judicature Society has concluded that the end product of the Judicial Commission has been good: "The Commission has succeeded in its mission of choosing candidates on the basis of professional competence."¹⁷⁵ Specifically focusing on the role of the lay participants, that study states "[s]urvey data permit the conclusion that lay contributions to the panels have been vigorous and useful. Just as a large number of lawyers may bring a variety of perspectives to the selection process, so may a sizable number of lay members. The perspectives of one group compliment the other."¹⁷⁶ The study concludes with the recommendation that the U.S. President should "appoint as many lay panelists as possible to the United States Circuit Judge Nominating Commission without doing violence to his wish to have each state represented by at least one lawyer."¹⁷⁷

An example of laypersons participating in an endorsement committee is provided by the Cincinnati Bar Association's Judicial Selection Committee. There, the bar association has a standing Judicial Selection Committee. While this Judicial Selection Committee conducts a poll of lawyers, it also solicits information from police officers, fellow judges, and others.¹⁷⁸ Ultimately, the Judicial Selection Committee weighs all of the information, including the bar poll, and rates each candidate in one of four categories: outstanding, well-qualified, qualified, and not qualified.¹⁷⁹ After an apparently successful experiment with non-voting lay person-

the lay participants in the judicial system. L. BERKSON & S. CARBON, *supra* note 27, at 163. In analyzing the data upon which this criticism was assessed, Berkson and Carbon indicate that "[b]ecause the criticism of nonlawyer participation was concentrated in three [of the thirteen] panels, it may be attributable to the performance of specific panelists rather than a flaw in the philosophy of allowing non-attorneys to participate in the selection process." *Id.* at 163 (footnote omitted). This criticism by panel members was also apparently echoed by at least three of the candidates for circuit judge that appeared before the United States Circuit Judge Nominating Commission. *Id.* at 170.

173. *Id.* at 158.

174. *Id.* at 191. Berkson and Carbon conclude that President Carter's program was an improvement, but that it was a merit selection for Democrats, rather than a non-partisan merit selection.

175. *Id.* at back cover.

176. *Id.* at 189.

177. *Id.*

178. C. PHILIP, *supra* note 58, at 25.

179. See generally Rueger, *Rueger Criticizes Judicial Selection Ratings*, CIN. B.A. REPORT, Dec. 1978, at 7. See also C. PHILIP, *supra* note 58, at 27. Several bar associations utilize a similar process. J. GUTERMAN & E. MEIDINGER, *supra* note 58, at 45.

nel on the Judicial Selection Committee,¹⁸⁰ the bar association amended its constitution to provide for the appointment of up to twenty-five percent non-lawyers as voting members.¹⁸¹ Though this proposition was adopted not without objection,¹⁸² it appears that the focus of the opposition seems to center upon the retention or selection goal of the Judicial Selection Committee that the "value of . . . professional opinion [will be] watered down if the final ratings and recommendations were directly influenced by opinions of non-lawyers."¹⁸³ The result was that as of October of 1979, the Cincinnati Bar Association president has appointed nine non-lawyers from the community to serve on the Judicial Selection Committee.¹⁸⁴

Though drawn from the selection retention context, these examples indicate that lay participation in the judicial performance context is possible and useful. Obviously, whether or not such

180. *Non-Lawyers Voted Full Voting Rights on Judicial Selection*, CIN. B.A. REPORT, Oct. 1979, at 7.

181. The reaction of the leadership of the Cincinnati Bar Association was generally favorable. It was said that former President John Getgey considered participation of the lay members of the committee to have been his greatest achievement during his tenure in office. *Judicial Ratings: Should Non-Lawyers Vote?* CIN. B.A. REPORT, Summer, 1979, at 6. The reaction elsewhere was described as generally favorable. *Rating the Non-Lawyers in Cincy*, CIN. B.A. REPORT, Nov., 1978, at 6. Though the Chairperson of the Judicial Selection Committee, Walter Bortz, later opposed giving laypersons the right to vote on the Committee, *Judicial Ratings: Should Non-Lawyers Vote?* *supra*, at 7, Bortz indicated that the lay members of the committee were "very diligent, very serious, very interested, very involved." *Rating the Non-Lawyers in Cincy*, *supra*, at 6.

182. The vote was opposed by a 522 to 360 margin. *Non-Lawyers Voted Full Voting Rights on Judicial Selection*, *supra* note 180, at 7. This constitutes an almost 60 percent vote in favor of lay voting.

183. *Judicial Ratings: Should Lawyers Vote?*, *supra*, note 181, at 7.

184. These individuals were as follows: Emile Godfrey, Jr., Operational Vice President of Federated Departmental Stores with responsibilities for monitoring legislation; John D. Geary, President, Midland Enterprises, parent company of Ohio River Company, also active in police justice affairs. Participants from 1979 included the following: R.A. (Bud) Anderegg, Executive Director, Cincinnati Business Committee and former Hamilton County Administrator; Lt. Donald I. Byrd, Commander, Homicide Section, Cincinnati Police Department; Lt. Ramon Hoffbauer, Squad Commander, Traffic Section and Auto Theft, Hamilton County Sheriff's Department; Dr. Jerome Jenkins, Executive Director, Seven Hills Neighborhood Houses, Research Associate and an adjunct professor; John Francis Kucia, Director, Treatment Alternatives to Street Crime, Cincinnati Health Department; Jane Juracek-Reherman, Manager, Public Affairs-Cincinnati Area for General Electric in Evendale; and Beverly H. Pace, Steering Committee member of the Junior League's Volunteers in Equity Program, currently investigating local mental health needs, former social worker in family counseling, also active in various volunteer community projects. CIN. B.A. REPORT at 7 (October 1979).

participation is desirable is a matter upon which disagreement can be had.¹⁸⁵ This is one issue which should be resolved in future study in this area.

Even after it is determined who will sponsor, formulate, and conduct the evaluation process, numerous other issues still remain to be resolved. Once the evaluation tool is devised, what process will be used to evaluate the evaluation? How can we check on the accuracy of the evaluation tool and whether it is doing its job?¹⁸⁶ Is it possible for the evaluation to be conducted by the entity charged with the oversight responsibility? Or is it necessary to have some third party actually physically conduct the evaluation process?¹⁸⁷ The same question as to objectivity should be explored with respect to where evaluation data is to be gathered or returned¹⁸⁸ and who is to compile and analyze such data.¹⁸⁹ The resolution of these issues is extremely important, since it will affect not only the validity of the evaluation process, but also the perception of the judiciary as to the reliability of the results.

C. *From What Sources Should Information be Gathered in the Evaluation Process?*

Traditionally the chief source of evaluative information on judges has been attorneys. This is, in part, because bar associations are the groups which have historically conducted most of the evaluations. In part, it may be the result of the fact that lawyers are the most cohesive and easily identifiable group with first-hand knowledge about the functioning of the judiciary. Also, the tendency to use attorneys to evaluate judges may be based upon the assumption that "[l]awyers as a group are particularly well qualified" to evaluate judges because "those attributes which col-

185. See, e.g., *Judicial Ratings: Should Non-Lawyers Vote?*, *supra* note 181, at 6-7.

186. "Results [of the Chicago Council of Lawyers Poll] of 1971 were tabulated, analyzed, and checked for internal and external validity by the Political Research Laboratory of the University of Wisconsin-Milwaukee." C. PHILIP, *supra* note 58, at 40. After the Maryland State Bar Association, *The Baltimore Sun* poll, a faculty member of Johns Hopkins University was hired to tabulate the responses and investigate their statistical validity. D. MADDI, *supra* note 51, at 63.

187. For example, in Arizona the questionnaires were sent out by and returned directly to the Survey Research Laboratory of the Department of Sociology of Arizona State University. D. MADDI, *supra* note 51, at 63.

188. In polls conducted by the Philadelphia Bar Association, the returns are sent directly to Price, Waterhouse and Company, for tabulation. C. PHILIP, *supra* note 58, at 22.

189. "The use of some sort of survey research organization, computer service, or accounting firm to tabulate the results often was seen both as an important safeguard to the integrity of the results and as an enhancement of the credibility of results disseminated to the public." D. MADDI, *supra* note 51, at 11.

lectively serve to outline the profile of a 'good' judge are ones which lawyers, as opposed to other groups or individuals, can both perceive accurately and evaluate meaningfully."¹⁹⁰

At the same time, the independent evaluations conducted by the media¹⁹¹ and those taken by the judiciary in New Jersey¹⁹² demonstrate that this is an area where new sources of evaluative information may be developed. "[T]he trend has been toward enlisting the participation of the wider community."¹⁹³

Obviously, the sources utilized for evaluation will vary depending upon purposes and uses of the evaluations and upon the "indicators of judicial performance" which are utilized.¹⁹⁴ For example, most of the sixteen objective indicators set forth for possible evaluation of trial judges can be gathered from court records.¹⁹⁵

On the other hand, more subjective information would come from individuals who had first-hand knowledge about a particular judge. The possibilities are numerous. For example, should the judges be required to evaluate themselves? Rather than having the judge respond to the evaluation initiated by others, it may be advisable to have the judge conduct a self-evaluation before receiving any outside information. This important idea was articulated by Justice Callow of the Wisconsin Supreme Court:

I have often thought that judges should be obliged to evaluate themselves and then let those who choose to disagree or support the evaluation come forward. Having the judge recognize that he is obliged to rate himself on a predetermined standard will encourage him to conduct himself within those standards so that he will not be subject to challenge.¹⁹⁶

A variation on this theme is found by the North Carolina Center for Public Policy Research in its evaluation of the judiciary of that state.¹⁹⁷ As part of an overall profile on the judges, the Center compiled information such as party background, surveys of the bar, and other similar information.¹⁹⁸ The Center also identified thirteen characteristics that might be identified with being a good

190. J. GUTERMAN & E. MEIDINGER, *supra* note 58, at 10.

191. Note 160 *supra*.

192. Handler, *supra* note 30.

193. C. PHILIP, *supra* note 58, at 1.

194. See generally notes 93-155 *supra* and accompanying text.

195. Note 149 *supra*.

196. Letter from Justice William G. Callow, Wisconsin Supreme Court, to Laurie O. Robinson, Director, A.B.A. Criminal Justice Section (March 4, 1980) at 2.

197. HARWELL, *supra* note 50.

198. *Id.* at 8-10.

judge.¹⁹⁹ In addition to having the bar rate judges on each of those characteristics,²⁰⁰ the Center asked the judges to identify and rank in order of importance the five characteristics of the thirteen that they considered most important for a judge to possess.²⁰¹

A second possible source of evaluative information is from other judges. Though peer review is not without its own controversies,²⁰² it is a tool which has been,²⁰³ and will continue to be utilized in the evaluation of the performance of lawyers.²⁰⁴ Currently information from other judges is utilized in at least two jurisdictions for evaluation, selection, or retention purposes.²⁰⁵

Even with respect to attorneys, the group from which evaluative data on judges is usually obtained, questions abound. Should information be sought only from attorneys who have actually appeared before a given judge? If so, should it be limited to those attorneys who appear regularly or who have made an appearance within a given number of years?

Another source of information, largely untapped, is laypeople who may have contact with the judge. Should some system be devised to attempt to obtain information from courtroom spectators? Should special efforts be made to have someone, whether a legally trained person or a non-legally trained person, act as a court observer with the specific purpose of evaluating a given judge's performance?²⁰⁶

Should efforts be made to obtain the input of jurors? This has been done in the past for use in retention decisions in Alaska,²⁰⁷ and juror information has also been successfully utilized in stud-

199. *Id.* at 235. These characteristics are as follows: common sense, professional integrity, knowledge of the law, open-mindedness, lack of prejudice, freedom from political influence, personal habits and conduct, hard working, diligent, native intelligence, intellectual honesty, basic understanding of human nature, courtesy to lawyers and witnesses, ability to make prompt decisions, and ability to keep control of case.

200. *Id.* at 229.

201. *Id.* at 235.

202. Compare Smith, *Peer Review: Its Time Has Come*, 66 A.B.A.J. 467 (1980) with Myers & Olrich, *Views of Our Readers*, 66 A.B.A.J. 820 (1980).

203. Dear, *supra* note 18.

204. A.B.A. Professional Competence *supra* note 25.

205. The Cincinnati Bar Association has used interviews of other judges in its evaluation. C. PHILIP, *supra* note 58, at 25. A variation of this idea is found in the New Jersey plan where appellate judges are to be given "important responsibilities in evaluating trial judges." Handler, *supra* note 30, at 5.

206. Though only two bar associations, The Association of the Bar of the City of New York and the Chicago Council of Lawyers, were identified as having a judicial observation program, it has been said that court watcher programs are increasing in popularity and effectiveness. C. PHILIP, *supra* note 58, at 45, 49.

207. The Alaska Judicial Council's evaluation included polls of jurors. Rubinstein, note 159 *supra*, at 480.

ies of felony cases involving evaluation of the performance of the prosecution, defense, and courts.²⁰⁸ The experience in the later study with respect to evaluative information obtained from witnesses, victims, and parties to the proceedings,²⁰⁹ suggests that these groups, too, should be considered as possible sources for information that might be useful in evaluating judicial performance.

Another group which regularly comes into contact with the court system and judges is law enforcement personnel. In the retention selection context, the Cincinnati Bar Association's evaluation committee seeks out the opinion of laypeople, including police officers.²¹⁰ Similar input is obtained in Alaska from the Alaska Peace Officers Association.²¹¹ It is clear that law enforcement personnel should be discussed in any plan considering the evaluation of judicial performance.

Should social scientists be asked to visit the courtroom to evaluate such things as a judge's speech or body language?²¹² Should performance evaluation include representatives of the media, who regularly cover court proceedings for the purpose of reporting the content of those proceedings? Should the media representatives who are specifically interested in evaluating the performance of judges be included? Should it include input from citizen groups who claim to maintain an active interest in the process of the judiciary? Should such a program make an effort to consult with

208. ANALYSIS AND DEMONSTRATION, *supra* note 140, at 399, 402. The data sought from the juror was largely subjective. Using a five-tier rating scale which ranged from agree strongly to disagree strongly, jurors were asked to rate four categories about judges and five categories related to the court system. "1. The judges were more interested in finding out the truth than in getting the case handled quickly. 2. The judges paid careful attention to the case. 3. The judges mainly favored the prosecutors and police. 4. The judges were willing to listen to both sides of the case." *Id.* at 361. "1. The court system is too slow and wastes a lot of time. 2. Things like race, family background, and the way a person looks make a difference in how a defendant or witness is treated in court. 3. The court system does a good job of protecting a defendant's rights. 4. The court system is too easy on defendants. 5. Many court officials and judges are dishonest." *Id.* at 363.

209. *Id.* at 400, 401, 403, 404.

210. C. PHILIP, *supra* note 206, at 25.

211. The Alaska Judicial Council's evaluation included polls of members of the Alaska Peace Officer Association. This organization includes not only voluntary membership by state and local police officers, but also correctional, probation and parole officers, and even federal agents from the FBI and Drug Enforcement Administration. Rubinstein, *supra* note 159, at 480.

212. For a recent account of the effect of a judge's body language and the improvements the judge can have by the "translation" of these "messages" see Givens, *The Way Others See Us*, 19 THE JUDGES, J. Summer, 1980, at 21.

people in the academic community who claim to keep track of the quality of the judiciary? Should the evaluation seek to elicit information from court personnel, such as court administrators, law clerks, staff members of the clerk's office, bailiffs and secretaries? Are there others who should be considered?

Each of these groups presents a possibly rich source of information that could be useful in evaluating judicial performance. The list is not meant to be exhaustive, nor is any group included for the purpose of claiming that that group's input is indispensable for a quality evaluation program. Rather, the objective has been to outline the alternatives and underscore the difficult choices that will be faced by anyone attempting to devise a thorough program of performance evaluation.

Indeed, those problems are magnified by concern with the possibility that some of the respondents within any given group will not be knowledgeable. The goal of eligibility criteria is to select a group of respondents representative of all knowledgeable respondents. The criteria should simultaneously maximize the inclusion of knowledgeable respondents and the exclusion of unknowledgeable respondents.²¹³

Even when dealing only with attorneys, polling practices have traditionally suffered from two potential defects. First, if information is sought from all attorneys the potential exists for a significant number of responses from those with no real first-hand knowledge. Second, attempts to limit those questioned to attorneys with first-hand knowledge significantly increase the cost of the evaluation process. Obviously, the inclusion of other groups, especially professional groups such as policemen and probation officers, could conceivably compound this problem.²¹⁴

D. How Should the Information be Collected?

One of the greatest methodology challenges lies in the collecting of the data necessary for a systematic multi-faceted evaluation.²¹⁵ To the extent that objective information is sought, it may be gained from first-hand inspection: checking dockets, reviewing case files, reading written opinions, and similar activities. In analogous situations students were used to compile this type of information, with apparent success.²¹⁶ Yet it must be recognized

213. D. MADDI, *supra* note 51, at 16.

214. One possible method of avoiding the problem of the unknowledgeable respondent would be to have the attorneys, jurors, policemen, and others fill out evaluation forms after *each* court appearance. This could, of course, cause tremendous administrative problems and be expensive.

215. CONCEPT PAPER, *supra* note 47, at 7.

216. *E.g.*, ANALYSIS AND DEMONSTRATION, *supra* note 140, at 276. For example,

that unless this data is the type normally recorded by the court systems for other purposes, its collection may be both labor-intensive and expensive.

An even more difficult problem lies with the collection of the subjective information from respondents who may have first-hand information about the performance of a given judge.²¹⁷ Any survey or poll of these individuals may well suffer from the defects upon which the traditional criticisms of bar polls are based: there could be a low response rate from those being surveyed and the lack of any controls insuring that those who do respond have adequate knowledge upon which to base an opinion.²¹⁸

In studies measuring the performance of certain phases of the criminal justice system,²¹⁹ a relatively high survey response was obtained from jurors,²²⁰ victims,²²¹ and other witnesses²²² by a procedure that involved mailing an initial survey, using a follow-up postcard, mailing a second copy of the survey, and if no response had been received by the foregoing efforts, making a follow-up phone call.²²³ Obviously one of the drawbacks of such an effort is that it is expensive.²²⁴

In the bar-polling context, attempts have been made to eliminate the unknowledgeable respondent by limiting those surveyed to attorneys whose names appear on the docket as having appeared before a judge within a certain specified period of time. In addition to the problems incidental to the fact that the docket may not accurately reflect all those who were actually involved in a given case or even who was the counsel who ultimately had primary responsibility for the case, the review of the dockets to retrieve the names of eligible respondents is, in and of itself, a demanding and costly project.²²⁵

Both of these problems may be surmounted by conducting per-

in the Multnomah County, Oregon project this data from court records was collected by a team of three law students, two court administration students, and a supervising staff member of the Rand Corporation.

217. See note 84 *supra*.

218. *E.g.*, D. MADDI, *supra* note 58, at 5-8; J. GUTERMAN & E. MEIDINGER, *supra* note 58, at 31-33; Marks, *supra* note 61.

219. Note 140 *supra*.

220. Note 216 *supra*, at 223.

221. Response rate from victims was 85%. *Id.* at 223.

222. Response rate from witnesses was 62%. *Id.*

223. Response rate from follow-up phone calls was 66%. *Id.*

224. *Id.*

225. Marks, *supra* note 61, at 38.

sonal interviews with a limited number of individuals who are thought to have sufficient knowledge of individual judges to make a useful assessment of the judge's performance. This is a technique that has been used in a limited sense in criminal justice performance evaluations²²⁶ and is frequently resorted to by journalists²²⁷ and advocated by those dissatisfied with the results of polls and surveys.²²⁸ This approach also raises significant questions. Of particular importance would be the issue of who is to choose the individuals to be interviewed. Another issue is the criteria to be utilized, and whether or not those interviewed will be given anonymity. Another alternative, set forth below, may lie in a system which would collect information from designated respondents immediately after each court appearance.²²⁹ This approach, however, also would suffer from substantial administrative and financial impediments.

There may be other alternatives, but those presented above should be sufficient to emphasize the extent of the problem in collecting the information needed for a useful evaluation of judicial performance. It is in this area that many of the most difficult problems lie and that substantial financial support will be necessary to employ social scientists to work with any group which evaluates judicial performance.

E. How Frequently Should Such Evaluations be Conducted?

Since polls and questionnaires have been traditionally used for retention or selection decisions, the frequency of their use has never been open to question. The poll would be taken whenever a new judge was to be selected or a decision was to be made on retaining an incumbent judge. Obviously, an evaluation program aimed at improvement of incumbent judges would not have its frequency pre-determined by election or appointment schedules.²³⁰ Thus, frequency of evaluation is a new issue. Actually, there may be two related questions. When should the data be collected? Also, when should it be processed and released to the judge being evaluated?

With respect to the first question, traditional thoughts would

226. Note 216 *supra*, at 424. This involved interviews of defendants in criminal proceedings.

227. *E.g.*, Atkinson & Stiteler, *supra* note 103; Pike & Crosby, *supra* note 160.

228. Flanders, *supra* note 66.

229. Collection immediately after court appearance could be done for every potential respondent within a designated group, *e.g.*, jurors, lawyers, or with a sample that was thought to be representative of the whole.

230. It has been suggested, however, that evaluations should *not* take place immediately before such selection/retention decisions. See J. GUTERMAN & E. MEIDINGER, *supra* note 58, at 52.

call for a single time period in which data could be collected, whether it be once a year, or once every five years, but other alternatives are available. For example, perhaps attorneys should fill out standard questionnaires immediately after each appearance before a judge.²³¹ Similarly, jurors could be required to complete such forms before being excused from jury service. Similar requirements could be made for witnesses, parties, and anyone else from whom information was sought. Thus, there could be an on-going collection of data.

Whether the data is collected at once, or in an on-going process, the question still remains as to when the evaluation is to be given to the judge in question. The only information on this question from other programs is that in New Jersey the recommendation is that no judge be evaluated until after two years of service; after that, all judges are to be evaluated every five years.²³² Obviously this issue is one which will have to be given sensitive consideration in any future inquiry.

F. Should the Raw Data Gathered Be Summarized, Refined, or Weighted?

One of the more important issues involved in evaluating judicial performance is in what form the data should be distributed. Should raw data be distributed? Should the data be summarized? Should it be refined or weighted in any way? Again borrowing from the experience with bar selection retention polls, it can be seen that all these approaches have been tried with varying success.

Some bar associations, conducting a multi-criteria poll with respect to a given judge, publish the full results for public consumption.²³³ Others do not reveal the statistical analysis behind the poll but do release to the public general summary information indicating whether a given judicial candidate is qualified.²³⁴ While both of these approaches are susceptible to the criticism that the results may be misinterpreted, misunderstood, or particularly in judicial election, misused, the level of controversy here is rela-

231. Letter from Professor Herbert S. Miller to Tom Karas (February 15, 1979).

232. Handler, *supra* note 30, at 10.

233. Santa Clara County Bar Association publishes such full results for the public. See generally D. MADDI, *supra* note 51, at 82-83.

234. Akron Bar Association Poll (unpublished poll on file with Pepperdine Law Review).

tively mild compared to that which has arisen under other approaches. Historically, an overwhelming number of bar polls have released raw data without any attempt to summarize or refine it.²³⁵ Yet in 1977, at least eight bar associations had some type of intermediary process by which poll results could be modified.²³⁶ In Cincinnati, Ohio, for example, "[t]he final rating decision is made by the Executive Committee of the Association which reviews the poll's findings. The Committee may modify them and only its ratings . . . are released."²³⁷

This process sparked public debate when an incumbent judge who had, in prior years been rated "well qualified," and who, in a later year received a plurality of "well qualified" votes on the bar poll, was nevertheless rated as only "qualified" by the Judicial Selection Committee.²³⁸ A subsequent president of the Cincinnati Bar Association indicated that the judge in question was "one of our most effective judges" and that he had been "deeply and personally offended" by the rating he received.²³⁹

While the Cincinnati Bar Judicial Selection Committee continues to "weigh" the results before it gives a recommendation to the public, this controversy apparently resulted in a change in the manner in which such a recommendation is given.²⁴⁰ In spite of problems such as these, there are strong advocates of using some committee or individual to weigh all the information and utilize a journalistic approach in summarizing the evaluation of a given judge.²⁴¹

A second example is provided by Alaska, where a decision was made to "weigh" the responses, not only as to category, but also as to the source. "[S]imply averaging the three ratings would make no sense. For example, 'sense of humor' could not be averaged on an equal basis with 'integrity.'"²⁴² Similarly, because jurors can often view a judge as one who can do no wrong because of what is termed as the "halo effect" and because they have the most limited exposure to the judges rated, their results were not given the same value as that of the attorneys and peace officers.²⁴³

235. J. GUTERMAN & E. MEIDINGER, *supra* note 58, at 27.

236. *Id.*

237. C. PHILIP, *supra* note 58, at 27.

238. *Rueger Criticizes Judicial Selection Ratings*, CIN. B.A. REPORT, Dec., 1978, at 7-8, with *Response of President*, CIN. B.A. REPORT, Oct., 1978, at 7-8.

239. *Newspapers Missed the Real Story on Judicial Selection*, McHenry says, CIN. B.A. REPORT, Oct., 1979, at 5-6.

240. *Id.*

241. *E.g.*, Flanders, *supra* note 66, at 304.

242. Rubinstein, *supra*, note 159, at 483.

243. *Id.*

It should be noted that this process was not without controversy. The Alaska Peace Officers' Association (APOA) gave a low, but "approval" rating, to the single judge who was recommended not to be re-elected once the overall evaluation was concluded. Though the APOA delegates apparently urged support of the recommendation, the APOA itself took an opposing position. It worked for successful retention of the judge in question, utilizing advertising that included the heading "Voters are you being misled?" and bumper stickers that read "Soft Judges Make Hardened Criminals."²⁴⁴

Since the purposes of evaluating judicial performance are different from the retention objectives of these programs, it is unlikely that an identical situation would arise. But these examples do illustrate the type of controversy that can result in situations where raw data may be refined.

Ultimately, the importance of the manner in which these issues are resolved may well depend upon the decision made as to the use of the data and its public distribution. On the one hand, if the information is to go solely to the judges evaluated, it would seem that the more information available the better. Raw data, summaries, and analysis could all be provided to the judges. Since the judges can make their own evaluation of the data, there would seem to be little harm from chances of misevaluation or misinterpretation. On the other hand, if the information is to become public, in whole or in part, accurate interpretation by third parties increases the seriousness of the issue and requires conscientious efforts at achieving an acceptable resolution of these questions.

VI. USES OF THE EVALUATION

The Evaluation of Judicial Performance Committee of the A.B.A. Criminal Justice Section has proffered an objective of self-improvement. If this objective is accepted,²⁴⁵ then the issues pertaining to the uses of the evaluation focus primarily upon the extent and form of the distribution of the results. Obviously, in order to improve judicial performance the results must necessar-

244. *Id.* at 481.

245. As noted previously, the author and the Evaluation of Judicial Performance Committee felt strongly that any proposed evaluation program should center upon on-the-bench improvement and avoid any attempt at effecting retention or discipline election decisions. Nevertheless, this is obviously an issue that any future inquiry committee could consider anew.

ily be available to the individual judge in question. Wider distribution presents serious questions. For example, in addition to evaluating their performance against an absolute standard, the judges may be interested in their performances relative to other judges within their own court or jurisdiction. Should they have access to such information? If so, should it be released with the name of each judge identified? Should the specific judges in question only be identified by a number or other code? Or should only a profile of the entire court be released? If some information is released under either of the three alternatives set forth above, how much should be released? Should it include detailed information or only summaries?

Beyond consideration of distribution to a judge's peers, should some form of the results be released to the judge's nominal superiors, such as an administrative judge or the state supreme court?²⁴⁶ Depending upon the form of such information, it might be helpful to these superior officials if the evaluation were to "(1) recommend areas of improvement . . . ; (2) create educational programs to redress the weakness; and (3) give the judges evaluated a heightened awareness and sensitivity to their own particular qualities and performance on the bench, thus providing a basis for an objective program of self-improvement and self-reliance."²⁴⁷

A similar, but perhaps more weighty issue is whether the results should be disseminated to the public in any form. On the one hand, it has been argued that any evaluation, even bar polls, should be released to the public because they will give incentive to the judges to improve. Since judges may be concerned with the public's view of them, public release may affect the judges' self-concepts in such a way as to lead to reassessment and improvement. Further, peer pressure may produce the same results. This seems to be the only theory behind the public release of bar surveys concerning sitting federal judges²⁴⁸ who are appointed for life and do not face an appointment or retention decision.²⁴⁹

On the other hand, strong objection has been made to the publication of such results. The theory of opponents of public release maintain that "any possible benefit could be gained . . . through

246. See, e.g., Handler, *supra* note 30.

247. CONCEPT PAPER, *supra* note 34, at 10.

248. E.g., Marks, *supra* note 61, at 37.

249. It has been suggested by one who opposes release of such data that the polls may be released simply to "publicize the activities of the bar associations." Flanders, *supra* note 66, at 304, 310. An additional reason sometimes given for public release is that it gives the bar a chance to publicly "blow off steam" when there is little else they can do with a judge with whom they are dissatisfied.

direct contact with the judge,"²⁵⁰ and that public release of the results may have detrimental effects since it tends to shift the emphasis from self-assessment and self-improvement to appointment or retention focuses. Particularly strong objection has been made to the release of such data shortly before an election in those states in which judges are chosen by direct ballot.²⁵¹

One interesting intermediate approach is that which the Arkansas Bar Association has taken with its bar poll, which includes an evaluation of the federal as well as the state bench. Under their plan, the results of the poll are given only to the judges. The provision was made that "[the] Judge may decide whether to release the results. If an intentional or significant error is made in the nature of the information released, the Association can release the entire survey for that judge."²⁵²

Of course, the fact that certain judges release their data and others do not might create public pressure upon others to do so, or might result in the media or public drawing negative inferences from those who do not make such a release.

The matter of public disclosure is more a question of judgment

250. *Id.* at 241. This approach was utilized by the Washington State Bar Association which, at least until 1977, did not make its bar poll results public. Rather, they shared the results with each judge individually. Marks, *supra* note 61, at 37.

251. Judge Burton S. Katz, of the Los Angeles Municipal Court has aptly pointed out some of the difficulties involved:

Any poll or questionnaire which is created for the purpose of evaluating judges *immediately* prior to an election would seem to be more susceptible to politicization and therefore entitled to little if any weight. . . .

I respectfully disagree with [the] suggestion that polls taken before an election would have the salutary effect of causing the judge 'to achieve higher quality judicial performance.' Such would not be the case, rather, the judge knowing that his/her future may depend upon the giving of 'popular' decisions or the posture of personal relationship with members of the bar or public may be inclined to compromise his/her sworn duty to objectively discharge judicial responsibilities without regard to personal consequences to the judge.

Letter from Judge Burton S. Katz, member of the A.B.A. Criminal Justice Section Judicial Performance Evaluation Committee, to Justice Rosalie E. Wahl of the Minnesota Supreme Court, then Chairperson of the Uses Sub-Committee and currently one of the Vice Chairpersons of the Evaluation of Judicial Performance Committee (November 1, 1979) (emphasis added). In a related matter it has been suggested that if the results are to be made public, then judges in states where the office is elected have a much higher stake in participating in the design of the evaluation program. CONCEPT PAPER, *supra* note 47 (letter from Chief Justice James Duke Cameron, Supreme Court of Arizona to Judge Jack Rosenberg (November 4, 1979)).

252. ARKANSAS BAR ASSOCIATION JUDICIAL POLL, JUDICIAL POLL 4 (1980).

than one of methodology. Nevertheless, it obviously calls forth strong disagreement and is an issue that any project on judicial evaluation must necessarily resolve.

VII. CONCLUSION

Evaluations of judges take place every day in a hundred different ways by thousands of different individuals. They are unavoidable because "[i]ndividuals inevitably judge one another in even their most superficial contacts, and they certainly do so in their daily work relationships. We all judge the quality of others' work, regardless of whether those judgments are rational, recorded, or expressed."²⁵³ This is also true of the judiciary. We have often been privileged to have the written opinions of jurors, attorneys, and parties about the quality of the judges before whom they appear. More recently we have experienced the phenomena of media or media-assisted evaluation of the judiciary.²⁵⁴ The evaluation by the media may be particularly unsatisfying, since media coverage is often superficial.²⁵⁵ Even when the evaluation is intensive, it may not produce any useful information which can be used for improvement.²⁵⁶ This should not be surprising, since the journalists' first objective is to attract the attention of the reading public, rather than to improve the judiciary.²⁵⁷

Thus, the Chairperson-elect of the National Conference of Special Court Judges, Judge James D. Rogers, observed that the evaluation of the performance of judges is inevitable.²⁵⁸ Judge Burton S. Katz found that such evaluations are likely to occur with "greater frequency and intensity."²⁵⁹ These observations are particularly apt. The question is not whether evaluation is going to

253. Jensen, *supra* note 35, at 36.

254. *E.g.*, *The Best and the Worst Federal Judges*, AM. LAW., July, 1980, 16-30. Within each federal circuit, a district judge was identified as the best judge and another as the worst judge. The criteria utilized was the following: "sheer legal ability, judicial temperament, willingness to work hard and integrity." *Id.* at 18. The data was gathered from the following sources: (1) "hundreds" of lawyers, (2) law professors, (3) local courthouse reporters, (4) prosecutors, (5) court of appeals judges, (6) opinions of the judges, and (7) opinions of the reviewing courts. *Id.* at 18. The final determination was made by the reporters: "it is journalism, not polling." *Id.* at 4.

255. See Griffin & Horan, *Judicial Merit Retention in Wyoming: An Analysis and Some Suggestions for Reform*, 15 LAND AND WATER L.J. 567 (1980).

256. *E.g.*, R. WOODWARD & S. ARMSTRONG, *supra* note 160. For this writer's view on this point, see Aynes, *Much Ado About Nothing*, 13 AKRON L.R. 507 (1980).

257. Interview with Bob Woodward and Scott Armstrong, authors of *The Brethren* at WGN-TV, Chicago, Illinois (December 10, 1979) at 13. "We're not looking for change. What we're doing it's [sic] *straight exposition*."

258. R. Kuh, Summary of Meeting on Evaluation of Judicial Performance, Honolulu, Hawaii, August 5, 1980 (August 18, 1980).

259. Letter from Judge Burton S. Katz, *supra* note 251, at 2.

take place, because it will. Nor is the question even whether results of such evaluations will be made known, because at least in some form, the media has let it be known that it is their intent to publish the results of their own evaluations of judicial performance. The only question is whether there should be a formalized, structured effort on behalf of the profession to utilize the evaluation that is already taking place to enhance the quality of the judiciary.

Obviously the Board of Governors of the American Bar Association, in authorizing the solicitation of funds to study a program of judicial performance evaluation²⁶⁰ and in authorizing the A.B.A. president to appoint a special seven-member committee to supervise such an inquiry,²⁶¹ has answered this question in the affirmative. The independent efforts of the New Jersey State Court system²⁶² and the National Center for State Courts²⁶³ in this same area demonstrate the timeliness and importance of action on proposals to evaluate judicial performance.

Though the A.B.A. Criminal Justice Section Evaluation of Judicial Performance Committee was re-appointed for the 1980-1981 year,²⁶⁴ its own work is largely completed.²⁶⁵ Hope for progress in formulating a positive and useful program for evaluating judicial performance rests on the ability of the A.B.A. president to utilize the proposed seven-member Special Committee on the Evaluation of the Judicial Performance to create a broadly based, repre-

260. Note 54 *supra*.

261. Note 53 *supra*.

262. Note 39 *supra*.

263. Note 40 *supra*.

264. All members of the Evaluation of Judicial Performance Committee, note 42 *supra*, were reappointed with Richard H. Kuh again serving as Chairperson. The Vice-Chairpersons for this year are Dean Robert B. Yegge, Justice Rosalie E. Wahl and the author of this article. New members appointed to the Committee include: Ephraim Gomberg, Criminal Justice Consultant, Palm Springs, California; Stephen Crane, Office of Court Administration, New York; Judge William R. Goldberg, Family Court, Rhode Island; Adrienne E. Volenik, National Center for Youth Law; Neil A. Kaplan, United States Department of Justice, Fraud Section, Criminal Division; Fred R. Harrell, Jr., Executive Director, North Carolina Center for Public Policy Research.

265. The goals of the Evaluation Committee for the upcoming year are the following: (1) to secure funding for the project; (2) to secure the necessary matching funds that might be required under any grant received; (3) to obtain the appointment of the seven-member Special Committee authorized by the A.B.A. Board of Governors at the Annual Meeting; and (4) to continue to gain the active involvement in the planning of this project of any other entities interested in evaluating judicial performance. Memorandum, *supra* note 53, at page 2.

sentative body which will work cooperatively with all interested parties in pursuing this project. Progress of this program also rests in the Law Enforcement Assistance Administration or some other funding source in providing the financial basis so necessary for the resolution of the complex issues involved in this undertaking.

POSTSCRIPT

Because of the major work that lies ahead in implementing a program to evaluate judicial performance, comments, criticisms, suggestions and information about other sources and resource persons are earnestly sought. This information will be gratefully received by Laurie O. Robinson, Director, Section of Criminal Justice, American Bar Association, 1800 M. Street, N.W., 2d Floor, Washington, D.C. 20036; or Professor Richard L. Aynes, University of Akron, School of Law, Akron, Ohio 44325.