

3-15-1997

## Evaluation of Administrative Law Judges: Premises, Means, and Ends

Ann Marshall Young

Follow this and additional works at: <http://digitalcommons.pepperdine.edu/naalj>



Part of the [Administrative Law Commons](#), and the [Judges Commons](#)

---

### Recommended Citation

Ann Marshall Young, *Evaluation of Administrative Law Judges: Premises, Means, and Ends*, 17 J. Nat'l Ass'n Admin. L. Judges. (1997)  
available at <http://digitalcommons.pepperdine.edu/naalj/vol17/iss1/1>

This Article is brought to you for free and open access by the School of Law at Pepperdine Digital Commons. It has been accepted for inclusion in Journal of the National Association of Administrative Law Judiciary by an authorized administrator of Pepperdine Digital Commons. For more information, please contact [Kevin.Miller3@pepperdine.edu](mailto:Kevin.Miller3@pepperdine.edu).

# EVALUATION OF ADMINISTRATIVE LAW JUDGES: PREMISES, MEANS, AND ENDS

A Proposal for Rethinking Traditional Models in Light of  
Due Process Concerns and Modern Management Theory and  
Research

**Ann Marshall Young\***

	Page
Introduction .....	3
I. Judicial Evaluation: The Current Context .....	10
II. Evaluation of Administrative Law Judges: Fundamental Premises .....	21
A. The basic principles of the American legal system should apply equally in the administrative legal system. ....	21
B. Independent decision-making is an indispensable part of the administrative adjudicatory system, and should be viewed less as a power than as a responsibility, which demands serious commitment on the part of decision-makers, and is not easily fulfilled. ....	24

---

\*Administrative Law Judge, Tennessee Department of State, Administrative Procedures Division, Suite 1700, James K. Polk Building, Nashville, TN 37243-0307. My thanks to Professor Harold Levinson of Vanderbilt University School of Law, and Edwin L. Felter, Jr., Chief Administrative Law Judge of the Colorado Division of Administrative Hearings, for their assistance and support in developing the ideas for this article; Professor Timothy Terrell of Emory University Law School, for reading an earlier draft of the article and sharing valuable suggestions on organization; Professor Canice Prendergast of the University of Chicago Graduate School of Business, for suggesting resources on the management principles discussed in the article; Judges Stanley L. Cygan, Manager of Administrative Hearings - Tax, Illinois Department of Employment Security, Edward J. Schoenbaum, also with the Illinois Department of Employment Security, and Blair Scoville Morgan and Mattielyn Williams, both of the Tennessee Department of State Administrative Procedures Division, for reading earlier drafts of the article and the original proposal for the fellowship paper, and providing helpful materials and feedback; and Patricia Hallan of Loyola University Chicago, School of Law, for her kind and patient assistance in preparing the article for publication.

- C. Management principles that promote the values of the American legal system and independent decision-making should guide the management of the administrative judiciary, including any evaluation, broadly defined as any standards-setting and enforcement in regard to ALJs. . . . . 31
  
- III. Traditional Evaluation Systems: Purposes and Unintended Consequences . . . . . 42
  - A. The purposes of traditional performance evaluation programs are generally good and well-intentioned, and include such things as productivity, efficiency, and quality and consistency of work product. . . . . 42
  - B. However, traditional evaluation programs, when used in an administrative adjudicative context, may have unintended negative consequences, including compromised independence in decision-making, lowered morale and quality of work, and decreased productivity. . . . . 43
  - C. The negative effects of traditional performance evaluation programs arise from the manifestations of human nature in an imperfect world and in imperfect systems; alternative approaches may more effectively achieve desired performance goals, foster independent decision-making, and fulfill the values of due process. . . . . 51
  
- IV. Proposal: Implementing Fundamental Principles in Traditional Settings . . . . . 54
  - A. Standards for Selection of Administrative Law Judges . . . 57
  - B. Standards of Conduct for Administrative Law Judges . . . 57
  - C. Continuing Education, Training, and Professional Development. . . . . 61
  - D. Ongoing Observation and Feedback on Performance . . . 63
  - E. Complaint and Disciplinary Procedures . . . . . 70
  
- V. Conclusion . . . . . 71

## Introduction

In today's climate of widespread public dissatisfaction with the American legal system, evaluation of judges has become quite popular. Seen as a means of improving the performance of judges, and of keeping judges accountable to the public, evaluation is touted as a solution of near be-all and end-all proportions.<sup>1</sup> Members of the public see evaluation of judges, including administrative law judges (ALJs), as the answer to the problem of an arrogant judiciary, unconcerned with the real problems of real people, and feel that through evaluations, they should be able to hold judges accountable, not only for their overall judicial conduct and legal ability, but also for the impact of their decisions in communities.<sup>2</sup>

---

<sup>1</sup> Northwestern University Professor of Law and Political Science Victor Rosenblum, in a report prepared in 1983 for the Administrative Conference of the United States, observed of evaluation generally that it "is, in some circles, the magic elixir that increases efficiency and effectiveness and identifies wasteful people and programs." Victor Rosenblum, *Evaluation of Administrative Law Judges: Aspects of Purpose, Policy, and Feasibility* 1 (Dec. 1983) (copy on file with the NAALJ Journal at Loyola University Chicago School of Law).

<sup>2</sup>One example of this is found in a Jan. 12, 1996, letter to the editor of THE TENNESSEAN, Nashville's morning newspaper, in which the writer complained of criminal offenders serving little time in prison and a third of violent crimes being committed by probationers and parolees, and asked, "Is not this a problem of judges and parole boards? Who judges their performance?" The writer continued:

A television expose told of militia groups setting up their own 'courts' with their own judges. This shows the frustration with the present system. Why can't we devise a method of performance evaluation of the results of the actions of our judges and parole board -- and an annual report to the public? Finally, we need a way to remove those who fail these standards.

Howard D. Meek, Letter to the Editor, THE TENNESSEAN, Jan. 12, 1996, at 12A.

United States House of Representatives Majority Whip Tom DeLay's call to impeach federal judges whose rulings are "particularly egregious," because "Congress has given up its responsibility in [overseeing] judges and their performance on the bench," appears to represent a similar anger at the judiciary and a wish to hold judges accountable for particular decisions that may be unpopular with a segment of the population. See Editorial, *Some Dangerous Talk About Federal Judges*, THE TENNESSEAN, Mar. 21, 1997, at 12A; Michael Kelly, TRB from Washington, "Judge Dread," The New Republic, March 31, 1997, at 6. Constitutional amendments ending lifetime federal judicial appointments or submitting them to re-confirmation every few years have also been proposed, and been criticized as leading to the "end [of] the independence of federal courts and opening] judges to charges of playing politics with rulings." Penny Bender, "Lawmakers Get Call to Impeach [U. S. District Court Judge

The legal community as well appears to have largely bought into evaluation of judges as a good idea, with a significant number of states and the Navaho Nation now having judicial branch evaluation systems either established or under development,<sup>3</sup> and with evaluation of state

---

*John T. J. Nixon*, "THE TENNESSEAN, Mar. 3, 1997, at 11A.

In response to the Tennessean editorial, which cautioned about the dangers of DeLay's proposal to the concept of an independent judiciary, the writer of another letter to the editor stated, "Independence is one thing, but arrogance is quite another. . . . Tom DeLay's method may appear to be madness to many scholars and elites. But the fact is that there is a tremendous desire from many people to do something positive about a judiciary system that has run amok." Steve Head, Letter to the Editor, THE TENNESSEAN, Mar. 24, 1997, at 12A.

<sup>3</sup>In 1993 there were 11 states (Alaska, Arizona, California, Colorado, Connecticut, Hawaii, Illinois, Maryland, New Hampshire, New Jersey, and Utah) and the Navaho Nation with established judicial evaluation systems, nine more states (Massachusetts, Minnesota, Missouri, New Mexico, North Dakota, Rhode Island, South Carolina, Vermont, and Washington) and Puerto Rico with systems under development, and one additional state (Delaware) with a system under consideration. *State Court Organization 1993*, by David B. Rottman, published by the Conference of State Court Administrators and the National Center for State Courts, Jan. 1995, NCJ-148346. As of 1995, Tennessee was developing a judicial evaluation program (which is now in effect), and Delaware, Maryland, and South Carolina were also in the process of developing programs. JUDICIAL PERFORMANCE EVALUATION HANDBOOK, 1996, American Bar Association National Conference of State Trial Judges.

Despite a recommendation in an August 1993 Report of the National Commission on Judicial Discipline and Removal, that judicial evaluation programs be adopted for the federal courts, the only form of evaluation under consideration at that time was self-evaluation by the judges, see James P. Timony, *Performance Evaluation of Federal Administrative Law Judges*, 7 ADMIN. L. J. 629, 655 at n. 142 (1993), and there are presently no official or formal judicial evaluation systems in any of the federal court districts or circuits. Interview with Deputy Public Information Officer Karen Redmond, of the Administrative Office of the United States Courts (Mar. 25 1997). However, biannual reports are required to be made on cases pending longer than specified time periods (from 6 months up to over three years), courts do informally monitor caseloads of judges, and some courts informally send questionnaires to lawyers. In addition, in the district courts, performance reviews of magistrates are done, and judges are aware of the views of lawyers on their performance through such publications as the *Legal Almanac*. *Id.*

With regard to Tennessee's system, an article in the July 1, 1996, Tennessean noted that just as the state appellate judges were about to institute a judicial evaluation system, "[l]ower-level judges [were] clamoring to be included in the sort of job appraisal that looks like the annual review now in use at many corporations." Larry Daughtrey, *Grade Us, Say Trial Judges, as Appeals Judges Will be*, THE TENNESSEAN, JULY 1, 1996, at 1B. The article continued, "Leaders of the low-key reform effort say their intention is to make the state's courts less intimidating and to give voters a better basis for an informed decision [on the retention of appellate judges] at the ballot box." The evaluations of trial court judges, who run in partisan elections, would not be made public and would be only for self-improvement. Tennessee's judicial evaluation program includes the production, by a judicial evaluation commission, of reports evaluating appellate judges based on information from judges, lawyers, and court personnel. These reports are made public prior to retention elections, in contrast to

ALJs, in various forms, being a reality in most states.<sup>4</sup>

Increased oversight and evaluation of ALJs have been advocated based on such justifications as improving ALJ skills and the quality of ALJ performance,<sup>5</sup> increasing ALJ efficiency and productivity,<sup>6</sup> combating bias and promoting consistency in ALJ decision-making,<sup>7</sup>

---

the confidential process whereby trial court judges select other judges to evaluate them for self-improvement purposes. As with other state judicial evaluation systems, responses by judges are permitted. Finally, judicial independence was one of the main concerns when the development of the program began three years ago, and the key to achieving self-improvement without compromising independence was considered to be selecting evaluation criteria that avoided any relationship to case outcomes. Interview with Harlan Goan, Assistant Director, Tennessee Judicial Performance & Evaluation Program, (Mar. 27 1997). See TENN. CODE ANN. §17-4-201 1994.

<sup>4</sup>April 7, 1997, Interview with Bruce T. Cooper, Administrative Law Judge with the Maryland Office of Administrative Hearings, and Chair of the NAALJ Judicial Evaluation Committee. According to Judge Cooper, this is especially the case in those states with central panels; other ALJs have suggested that evaluation of ALJs is found more often in agency ALJ offices. The evaluation committees of the NAALJ and the ABA National Conference of Administrative Law Judges are presently, on a joint basis, surveying and analyzing the practices of the various states, but there are no actual statistics on this as yet. It is noted that some federal administrative judges and hearing officers may also be evaluated; however, evaluation by agency personnel of federal administrative law judges appointed pursuant to 5 U.S.C. §3105 is not permitted, see note 32, *infra*, and attempts to establish such evaluation systems have been opposed by ALJs, on the grounds that this would impede judicial independence. See, e. g., Timony, *Performance Evaluation of Federal Administrative Law Judges*, 7 ADMIN. L. J. 629 (1993); see also note 32.

<sup>5</sup>See, e.g., Rosenblum, *supra*, note 1, at 72-73. The former Administrative Conference of the United States (ACUS) in its Recommendation 92-7, 1 C.F.R. §305.92-7, 57 F. Reg. 250, Dec. 29, 1992, concerning the Federal Administrative Judiciary, would take the goal of improving ALJ performance to the extent of having chief ALJs evaluate ALJs on their adherence or nonadherence to agency policy. Richard Sippel's article in this issue of the NAALJ Journal covers this subject very well, and, other than to agree that such evaluation could compromise judicial independence, ethics, and collegiality, this author will not specifically address the contents of Recommendation 92-7 in this regard. See also Ronnie A. Yoder, *Evaluation: Where Are We? Where Are We Going?* 14 J. NAALJ 303, 309 (1994), in which Yoder observes that "evaluations, particularly when tied to 'compliance with agency policy' and separated from 'good cause' discipline before an impartial adjudicator is an invitation to agency coercion -- subtle or not-so-subtle -- of ALJ independence."

<sup>6</sup>See Jeffrey S. Lubbers, *The Federal Administrative Judiciary: Establishing an Appropriate System of Performance Evaluation for ALJs*, 7 ADMIN. L.J. AM. U. 589 (1994).

<sup>7</sup>See ACUS Recommendation 92-7, *supra* n. 5. ("[A]gencies have a legitimate interest in being able to manage their employees, including ALJs, in order to ensure that the adjudicatory system is an efficient and fair one.") See also Elaine Golin, *Solving the Problem of Gender and Racial Bias in Administrative Adjudication*, 95 COLUM. L. REV. 1532, 1544 (1995) (addressing statistical analyses that have shown racial disparities in Social Security hearing outcomes).

Columbia Law Professor Richard Pierce has a somewhat different perspective on the

---

issue of bias. Pierce argues that “agency policy should reflect bias -- the bias of Congress or, in the absence of legislative expression of that bias, the bias of the President.” Pierce, *Political Control Versus Impermissible Bias in Agency Decisionmaking: Lessons from Chevron and Mistretta*, 57 U. CHI. L. REV. 481, 486 (1990); see also KENNETH CULP DAVIS AND RICHARD J. PIERCE, JR., ADMINISTRATIVE LAW TREATISE, 3d. Ed., §9.10, at 103. The type of bias Pierce contemplates in this statement is obviously quite different from the sort of bias Golin sees as a problem. Indeed, he calls consistency in decision-making an important due process value. *Supra* Pierce, at 512. On the appropriate methods to achieve such consistency, Pierce argues that the Social Security Administration’s efforts in the 1980s to mandate the proportions of decisions granting and denying benefits were “entirely proper,” and asserts that the Supreme Court would uphold such methods and reverse lower court decisions that have struck such efforts down. *Id.* at 483-4.

Pierce’s perspective on the extent of oversight of judicial decision-making that would be proper is perhaps best illustrated by a passage from his article in which he raises several questions, to demonstrate his theory that federal judges are, inappropriately, unwilling to allow politicians to make policy decisions, because of a combined “ignorance of alternative decisionmaking procedures [and] a distinct bias rooted in their role in government.” *Id.* at 516. Pierce on federal judges:

Judges control the judicial decisionmaking process. The identification of weaknesses in that process, or alternatives that perform better in some contexts, threatens their self-image. Many are unwilling to admit that, in some contexts, judicial decisionmaking is absurdly expensive, highly subjective, and rife with inconsistency. Moreover, initiatives like the SSA’s efforts to control the conduct of its ALJs strike far too close to federal judges’ own turf. What if some institution actually began to monitor the productivity and inter-judge consistency of the federal bench? What if such an investigation detected major differences in productivity and large inconsistencies in outcomes? If the SSA can exercise control over the productivity and consistency of its ALJs, perhaps some institution has the power to exercise analogous control over federal judges. In short, federal judges are biased decisionmakers when they draw lines between permissible political control of agency policymaking and impermissible bias in agency decisionmaking.

*Id.* at 516-517. Pierce recognizes in a footnote that “[s]tatistically-based measures, analogous to the SSA’s benchmarks for ALJs, would be a poor method of controlling federal judges. Statistical measures of performance would be valid only if judges were randomly assigned large numbers of relatively homogenous cases.” *Id.*

*But see* Timony, Performance Evaluation, *supra* note 3, at 639, in which Judge Timony points out that there are certain types of cases heard by federal judges that seem to have as many similarities as Social Security disability cases. (“Prisoner petitions, bankruptcy petitions, and run-of-the-mill criminal and civil cases seem to have as many similarities as disability cases.”). *Id.*

Without trying to compete with my law school professor who once wrote an article with a seven-page footnote, I would question whether Pierce’s analysis omits a crucial inquiry: How will the performance of the investigators, monitors, and evaluators of judicial performance be monitored, evaluated, and judged, and who will be the last-resort, final judge/evaluator? The professor seems to believe that there is “some institution” somewhere that has the capability to reach appropriate judgments about the performance of judges that either could or should not

and encouraging positive reinforcement for ALJs who perform well.<sup>8</sup> All of these are worthy ends. However, it is the contention of this article that, unless sufficient attention is paid to the means used to achieve such ends, and to the need to protect judicial independence<sup>9</sup> on

---

be questioned (and despite the question form used and the limiting footnote, he appears to believe this with regard to both ALJs and judicial branch judges). He further seems to believe (with regard to statistical information about ALJ cases and decisions) that statistics are never misleading or manipulable, and that the final arbiter of questions of judicial performance should be a political entity or person. His analysis in effect questions many of the fundamental premises of our legal and governmental system, including the recognition that, although no form of decision-making will ever be perfect, there is a need for some last-resort process for deciding disputes; that the procedures generally encompassed in the judicial decision-making model -- refined continuously over time through both the common law and statutory law-- are what this society has arrived at as the best such last-resort process; that this process has, at least to the present date, held up in our system of checks and balances; and that an evidentiary hearing, with parties present, with notice and opportunity to respond to allegations, with a neutral decision-maker(s) required to make a decision based on the record produced at such hearing, and with the right of appeal, is an integral part of such a last-resort process.

Although it is no doubt true that there are biased judges, it is questionable whether more oversight of judges of the sort Professor Pierce advocates will, on balance, reduce the incidence of such bias overall, given the danger of chilling independent and impartial decision-making that can result from such oversight. See CHARLES H. KOCH, JR., *ADMINISTRATIVE LAW AND PRACTICE*, §6.7[2] at 218-219 (1985, Pocket Part 1997); *Barry v. Bowen*, 825 F.2d 1324, 1330 (9th Cir. 1987), *appeal after remand* 884 F.2d 442 (9th Cir. 1989) and cases cited therein. Judicial independence is protected, and judges are required to comply with requirements that place stringent limitations on their activities in general, *in order to assure* that they are as neutral, impartial, fair and unbiased as possible in their decisions. (Other aspects of the legal system as well serve to assure that this goal is achieved: the adversary system of justice, rules of evidence, etc.)

Certainly, as discussed in the text of the article, judicial branch judges and ALJs should monitor and police themselves, appropriately, and where necessary, standards of conduct and enforcement of such standards through complaint and disciplinary procedures should be strengthened. However, to allow the sort of oversight Professor Pierce advocates would in effect substitute, in the place of a hearing conducted pursuant to principles of due process, a new last-resort decision-making process at the evidentiary hearing level. Such a process would have no built-in protections such as those that have been developed over time in the judicial model. Moreover, any such oversight process would itself -- to follow Pierce's argument to its logical conclusion -- require some oversight from yet another institution, to assure that *it* is consistent, fair, appropriate, etc. The image of looking into a mirror facing another mirror and seeing an infinite number of images receding into the background comes to mind.

<sup>8</sup>KOCH, *supra* note 7, §6.3 at 207 (Pocket Part 1997).

<sup>9</sup>See Rosenblum, *supra* note 1, at 27: "Because judges are so often thrust into bitter social controversies, the 'first principle' in any judicial evaluation system is the obligation to preserve the independence necessary for judges to perform properly the judicial task." Rosenblum goes on to observe:

No well-structured judiciary can disregard the fundamental value of public



a practical and human basis, the costs of such oversight and evaluation may outweigh any potential benefits.

It is therefore of crucial importance to analyze closely, in light of available knowledge, the full range of possible unintended consequences of any proposed evaluation standards and procedures. In this regard, modern management research and theory suggest that traditional performance evaluation, based on management by objectives and the grading or ranking of those evaluated on point scales, can produce unintended negative consequences that may compromise the integrity of administrative adjudicative systems.

Indeed, the more that effective job performance in any field

---

accountability in a democracy; and everyone involved in and affected by the judicial system deserves assurance that the system can and will respond to allegations of unfitness . . . . A featured place on judicial evaluation agendas must be accorded the issue of how best to identify 'the threshold of pettiness' that can achieve the best possible accommodation between principles of judicial independence and public accountability.

*Id.* at 28. See *infra* note 123, on the particulars of Rosenblum's suggested evaluation program for ALJs.

See also Diana Farthing-Capowich, *Designing Programs to Evaluate Judicial Performance*, 9 ST. CT. J. 22, 23 (1985):

[W]here the product is justice, both output and the quality of the product are difficult to delineate, monitor, and assess. Furthermore, the legal and philosophical underpinnings of the judicial system encompass the broader concepts of judicial independence and social organization and stability. Evaluation of the judiciary must not encroach upon judges' ability to render decisions in a truly impartial fashion (e.g., the scope of evaluation programs should not include ideological review).

See also L. Hope O'Keeffe, *Administrative Law Judges, Performance Evaluation, and Production Standards: Judicial Independence versus Employee Accountability*, 54 GEO. WASH. L. REV. 591, 626 (1986):

At a minimum, a properly designed evaluation program must protect judicial independence, guarantee procedural fairness and substantive impartiality in the program's design and implementation, and accord due consideration to the quality as well as the quantity of ALJs' decisions. Anything less might increase the accountability of ALJs at the risk of imperiling their independence and the integrity of the APA's system of adjudication. (Citations omitted.)

The ways in which evaluation can compromise independent thinking, in more subtle ways than many have perhaps considered, has been addressed in management literature, which is discussed in the text accompanying notes 61-86. The ways in which independent and impartial decision-making can be compromised has also been addressed in the legal literature. See text accompanying notes 40-47. See also, *infra*, discussion of unintended consequences of performance evaluation.

involves independent decision-making, the less effective traditional performance evaluation may be in producing quality, and quantity, in performance. According to recent innovative management studies, such evaluation and other ranking mechanisms can tend to create "yes" men and women, compromise productivity and morale, decrease efficiency, and be corrupted by bias and favoritism.<sup>10</sup> This is particularly undesirable in a system that is held out to the public as assuring neutral and fair decision-making on important matters in people's lives.

The first section of this article consists of a review of the current context of judicial evaluation. An understanding of the setting in which most calls for judicial evaluation arise and in which evaluation programs are implemented is helpful, in order to appreciate fully the possible effects and consequences of such evaluation. Without such a real-world perspective on judicial evaluation, any analysis will likely be purely theoretical and of little value. Careful examination of the real-world context for such evaluation provides a basis, both for a clearer view of the subject in light of relevant legal and management principles, and for more practical and effective application of appropriate principles.

In Section II of the article, I propose that any programs to evaluate administrative law judges should be grounded in three fundamental values or premises: A. that the basic principles of the American legal system -- as neutral as possible application of case precedent and lawfully enacted and promulgated statutes and rules, rather than justice according to personal or political viewpoints of what is right -- should equally apply in the administrative legal system; B. that independent decision-making -- an integral part of the American legal system -- is likewise an indispensable part of the administrative legal system, and should be viewed less as a power than as a responsibility, which demands serious commitment -- and sometimes courage<sup>11</sup> -- on the part of ALJs, and is not easily fulfilled; and C. that sound, effective management principles that promote the values of the American legal system and independent decision-making should guide the management

---

<sup>10</sup>See *infra* text accompanying notes 61-86.

<sup>11</sup>See Penny J. White, *Judicial Courage and Judicial Independence*, XVI J. NAALJ 161 (Winter 1996).

of administrative adjudicative systems, including any evaluation of ALJs -- defined broadly as any standards-setting and enforcement with regard to ALJs.

With these premises, and with the assistance of some research and theory on performance evaluation from the field of management and economics, I will then, in Section III, consider the purposes of evaluation, and address some of the ways in which traditional ALJ evaluation systems can have unintended consequences that may negatively affect the performance of ALJs in fulfilling their central duty to decide cases independently, impartially, and fairly.

Section IV of the article contains suggestions for crafting a system -- either from the ground up or as a modification of an existing system -- that will best promote the responsibility of independent decision-making, and most effectively enhance the quality, efficiency, productivity, and fairness of ALJ performance.

Finally, in the conclusion of the article, I propose that creating standards-setting and enforcement systems that take into account the basic principles of the American legal system, and what we can learn from the field of management, should both improve ALJ performance and encourage ALJ fulfillment of the responsibility of independent decision-making, and should thereby more effectively and appropriately address public discontent with the legal system, through reassurance that the system is in fact fair, as it is intended to be.

## I. Judicial Evaluation: The Current Context

Both judicial branch judges and administrative law judges function in the real world of public opinion, politics, and public budgets. Notwithstanding efforts to insulate judges from political and other influences, these realities are always there, in the background, capable of exerting sometimes very real pressures on both federal and state judges and ALJs at all levels, whether for good or bad depending upon the point of view.<sup>12</sup> The present relationship between the legal/judicial system and the public may, at the very least, be described as one containing elements of tension: between respect and less temperate

---

<sup>12</sup>See *supra* note 2 and accompanying text; see also *infra* note 93.

points of view, and between ideas of judicial accountability and judicial independence. These elements of tension play a significant role in discussions of judicial evaluation.

The importance of maintaining judicial independence is generally acknowledged and emphasized among those in the legal community, and indeed today is a subject that is given much attention, with the last issues of both *Judicature* and *The Judges' Journal* devoted to this subject.<sup>13</sup> However, except in the context of the administrative judiciary,<sup>14</sup> the fundamental propriety of the evaluation of judges as it

---

<sup>13</sup>See 80 *JUDICATURE* (Jan.-Feb. 1997); 36 *JUDGES' JOURNAL* (Winter 1997). It is noted that in some quarters there is no presumption of independence in ALJ decision-making; for example, it has been observed that "the employer-employee relationship between agencies and ALJ's gives rise to a public perception that ALJ's are not unbiased or impartial judges." Karen S. Lewis, *Administrative Law Judges and the Code of Judicial Conduct: A Need for Regulated Ethics*, 94 *DICK. L. REV.* 929, 930-931 (Summer 1990). Indeed, there are those who argue in favor of limiting ALJ independence. See, for example, *supra* note 7. However, it is generally agreed that ALJs should be independent adjudicators, in order to satisfy basic values of due process. See *infra* text accompanying note 51. The more relevant debate usually centers on how, as a practical matter, to balance judicial independence with accountability. See Edwin L. Felter, Jr., *Maintaining the Balance Between Judicial Independence and Accountability in Administrative Law*, in this issue of the NAALJ Journal.

<sup>14</sup>There is a significant body of literature on the propriety of evaluating and otherwise overseeing the work of ALJs, primarily concerning federal ALJs, primarily Social Security ALJs. This article refers to some of this, but does not purport to exhaustively address such sources. This article is intended to approach the issues from a broader perspective, encompassing state as well as federal ALJs, and considering management as well as legal sources. Since the author's experience is as a state ALJ, the article may reflect this perspective; however, the principles addressed are intended to have more general application.

It is noted that, in this article, the term administrative law judge or ALJ will be used, and is intended to include all those who decide administrative law cases after holding evidentiary hearings in which records are created and which are governed by administrative procedures acts or similar law or regulation, whether they be called administrative law judge, administrative judge, hearing officer, hearing examiner, or any other appellation. Also, while the author is a lawyer and the article may reflect this in sometimes appearing to assume that ALJs are always lawyers, it is recognized that many ALJs (whatever they are called) are not lawyers, and that non-lawyer ALJs can be good ALJs. However, the author strongly believes that all the legal principles discussed in the article apply equally to non-lawyer ALJs, since as ALJs they are performing a legal function subject to the same fundamental requirements of due process and fairness that govern the actions of any ALJ. They should also be held to the same standards of judicial courage of which former Tennessee Supreme Court Justice Penny J. White so eloquently spoke at the November, 1996, Annual Meeting and Conference of the NAALJ. See note 11, *supra*. It is furthermore recognized that lawyers have no special abilities in this regard, notwithstanding that the law can help us to define more clearly when and how it may be necessary to exercise courage, and can be a strong support when such exercise becomes very difficult to undertake or sustain.

relates to judicial independence has not been extensively disputed in the literature, to the knowledge of the author.<sup>15</sup>

Commissions and committees studying and/or implementing judicial evaluation systems in the court system see judicial independence as a central concern, in the context of arriving at appropriate criteria and procedures for evaluating judges,<sup>16</sup> but appear generally to conclude that the benefits of judicial evaluation outweigh the dangers of compromising judicial independence, or that any such compromise can be avoided through careful definition of evaluation criteria. Indeed, it has been stated that performance evaluations of judges, in addition to providing valuable feedback to judges on how well they are carrying out their duties, "also preserve judicial independence by focusing the electorate away from individual judicial decisions and more properly toward the judge's overall performance as

---

<sup>15</sup>Judge Irving Kaufman's observation, that "statutes that vest independent entities, groups of judges, or private individuals with authority to reward or punish judges based on their performance" were "questionable," is noted. Irving R. Kaufman, *The Essence of Judicial Independence*, COLUM. L. REV. 671, 696 (1980) (citations omitted). Also noted is Judge Randall R. Rader's recognition that "[f]or both federal and state judges, the major obstacle to attorney evaluation is the perception of a threat to judicial independence." U. S. Court of Appeals Judge Randall R. Rader (D. C. Circuit), *Evaluate Your Own Performance on the Bench*, 30 JUDGES' JOURNAL 32, 34 (Summer 1991). Judge Rader considered such resistance to evaluation on the part of judges to be "natural": "Judges, after all, must enforce the law without regard to perceptions, criticisms, or public opinion, [and] remember well, when they were practicing lawyers, the spontaneous, unflattering evaluations they gave to judges when they were ruled against." He concluded that "[o]ne overriding consideration . . . counsels judges to seek evaluation. Judges control lives and destinies in their decisions. Therefore, a judge has a professional responsibility to seek every credible avenue of self-improvement." *Id.* Judge Rader observed that "[t]he real question is how to solicit feedback without harming judicial independence," and suggested the necessity of a confidential and candid atmosphere, and the use of evaluations for self-improvement. *Id.*

See also Richard Aynes, "Evaluation of Judicial Performance: A Tool for Self-Improvement," 8 PEPP. L. REV. 255, 311 (1981). Aynes likewise advocated the use of evaluation for self-improvement, observing that evaluation of the performance of judges has been called "inevitable," because it is already taking place "every day in a hundred different ways by thousands of different individuals," including the media, in "often superficial" ways. "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."

Despite such sources as these, there appears to be little literature on the subject of the dangers to judicial independence that may be involved in non-confidential evaluation of judges, in which judges may be rated, and which may have an effect on the job security of judges.

<sup>16</sup>See, e.g., interview with Harlan Goan, *supra*, note 3.

a public servant.”<sup>17</sup>

Nonetheless, while judicial evaluation commissions, and the judicial and legal community in general, may be cognizant of the need to avoid compromising judicial independence, the general public is less likely to be. Reacting with fear, anger, and impatience to their own perceptions of problems that have caused them or others known to them real difficulty and pain, citizens express their concerns with an emotional urgency that tends to elicit popular sympathy and drive public opinion.<sup>18</sup> Whereas in the past the role of the judge was

---

<sup>17</sup>Editorial on judicial independence, 80 JUDICATURE 152 (1997). Addressing the issue of concrete means by which an independent judiciary can be maintained, the editorial continues:

While judges must be free to use their best legal judgment in reaching decisions regardless of how unpopular those decisions may be, in no way does judicial independence give carte blanche to judges to violate the code of judicial conduct or otherwise fail to perform their duties responsibly. Over the years [the American Judicature Society] has recognized this through activity in two areas: judicial performance evaluations and judicial conduct and ethics.

Observing that “a truly independent judiciary is possible only when the public accepts the legitimacy of judicial authority,” the editorial notes the present apparent erosion of public confidence in the judiciary, calls for judges to reach out in educational forums outside the courtroom, and concludes that, “[i]f the role of the courts in our democracy is not thoroughly explained to the public, the public is susceptible to misleading or inaccurate information that threatens the maintenance of an independent third branch of government.” See also Frances Kahn Zemans, *From Chambers to Community*, 80 JUDICATURE 62 (Sept.-Oct. 1996).

It may be observed that, if the judicial branch of government is having a difficult time explaining its role to the public, the administrative judiciary’s ability to explain its role in a way that is understandable to the public (and indeed, to other governmental officials) may be said to be even more problematic. In this vein, the relative lack of prestige of ALJs may, ironically, stand us in good stead with a public that is suspicious of judges it perceives to be “arrogant.” How to maintain respect for the integrity of the proceedings we conduct, while at the same time avoiding the appearance of being arrogant – in the face of the perception by many ALJs of a dismissive inclination toward the ALJ role (and even the use of the term “judge”) on the part of some in the legal, governmental, and academic communities – is a challenge ALJs must continually address.

<sup>18</sup>On the power of public opinion, as early as 1835 Alexis De Tocqueville predicted that “faith in public opinion will become . . . a species of religion [in the United States], and the majority its ministering prophet.” ALEXIS DE TOCQUEVILLE, *DEMOCRACY IN AMERICA*, 1840, THE HENRY REEVE TEXT, Part II, First Book at 11. (Phillips Bradley ed., Alfred A. Knopf, 1945). Further, more ominously:

[I]n a democratic country . . . public favor seems as necessary as the air we breathe, and to live at variance with the multitude is, as it were, not to live. The multitude require no laws to coerce those who think not like themselves: public disapprobation is enough; a sense of their

considered to be unquestionably respectable if not sacrosanct,<sup>19</sup> in present times harsh attacks on judges and other public figures are common.<sup>20</sup> And just as public outrage has manifested itself in attempts to oust judges whose decisions do not please critical quotients of the voting populace,<sup>21</sup> it has surely played a role in the growing popularity

---

loneliness and impotence overtakes them and drives them to despair.

. . . . In whatever way the powers of a democratic community may be organized and balanced, it will always be extremely difficult to believe what the bulk of the people reject or to profess what they condemn.

*Id.*, Part II, Third Book, at 261. And moreover, in De Tocqueville's view, "the government likes what the citizens like, and naturally hates what they hate." *Id.* at 295.

<sup>19</sup>See *infra* note 39.

<sup>20</sup>It may fairly be observed that those of us who are lawyers may share responsibility for the low levels to which public discourse in and about the legal system has sunk. ABA President N. Lee Cooper has called for a reversal of the "decline of civility" in our justice system. *President's Message*, 83 A.B.A.J. 8 (March 1997). See also Marvin E. Aspen, *What We Can Do About the Erosion of Civility in Litigation*, 35 JUDGES' JOURNAL 32, 35 (Fall 1996), in which leadership on the part of judges is called for to promote civility in the practice of law and "to renew respect for, and confidence in, the judicial system"; and Charles Mahtesian, *Supreme Chaos*, GOVERNING, (July 1996) at 40, in which allegations of conspiracies, phone-tapping, case-fixing, sexual harassment, secret slush funds, wrestling matches between judges, and payoffs on the part of State Supreme Court Judges, are described.

Whether the strong feelings that seem to underlie such uncivil interaction have simply moved from parts of society where they always existed to quarters that were previously more genteel is uncertain. Or perhaps it was always there but was just not reported, as Mahtesian points out. *Id.* at 41. It has been observed that some of the political invective to be heard and read in the early years of our democracy rivaled, if not surpassed, that heard today. See Stephen B. Burbank, *The Past and Present of Judicial Independence*, 80 JUDICATURE 117, 121 (Jan-Feb. 1997). It is also possible that the eruption of extreme sentiments in the arena of law and public policy may be some sort of cyclical phenomenon, or the result of forces beyond our ordinary perceptive powers -- the equivalent, for example, of the fluttering of a butterfly's wings that starts a movement of air that ends up as a tornado a continent away. JAMES GLEICK, *CHAOS - MAKING A NEW SCIENCE*, 8, 322, n. 20 (1987). In any event, it seems clear that much of the legal system perceives itself to be under attack (whether for valid or invalid reasons is beyond the scope of this article), and the establishment of judicial evaluation systems appears to be at least partially a defensive maneuver, even if self-improvement is part of the motivation for the creation of such systems. See also *supra*. note 15.

<sup>21</sup>See, e.g., John Gibeaut, *Taking Aim*, 82 A.B.A. J. 50 (Nov. 1996), which begins, "To some, the rash of political attacks on judges is an assault on the very concept of judicial independence. To others, it is merely robust debate. Whatever the interpretation, the personal tone of attacks on isolated decisions is shaking up the system." *Id.* The article describes attacks on several judges, including Justice Penny J. White of the Tennessee Supreme Court, who in August 1996 lost a retention vote after various groups campaigned against her, largely on the basis of her vote in a single death penalty case, remanding the case for a new sentencing hearing. Tennessee's evaluation program for judges in retention elections was not yet in place at the time. Whether a favorable evaluation would have prevented the vote against Justice White is an unanswerable question; future elections may provide insight in this regard.

of evaluation, seen by some as a means of controlling judges as perceived agents of the public.<sup>22</sup>

During such times of widespread public loss of faith in and respect for formerly revered institutions, society goes through varying degrees of upheaval, with factions at odds with each other on multiple, overlapping fronts. This has been so in the administrative legal system in recent years, and may be so to an even greater degree in the near future. Change is proposed, or objected to, on the basis of grounds that seem eminently reasonable to their champions, but time, political and other pressures limit effective examination of all the potential consequences of proceeding in one way or another. Public colloquy often seems to follow the image of a pendulum that swings from one end of a spectrum of possible viewpoints on an issue, to the other.<sup>23</sup>

The administrative adjudicatory system may be particularly subject to attempts, by parties from all across the political/public policy spectrum, to modify or modulate it. The prevalence of such efforts is likely related to the status of the administrative law system as a statutory and regulatory creation with no constitutional mandate or

---

<sup>22</sup>See *supra* note 2. There is little doubt that, whether or not articulated as such, the view that all public officials including judges are not only servants, but also agents, of the public, plays largely into much public discourse about evaluation of judges, sometimes with little appreciation being shown for the subtleties of what constitutes the appropriate duties of officials in all three branches of government. Professor Pierce has written on his agency theory of government, arguing for more deference on the part of courts to agency policy-making when reviewing the actions of administrative agencies. Richard J. Pierce, Jr., *The Role of the Judiciary in Implementing an Agency Theory of Government*, 64 N.Y.U. L. REV. 1239 (1989). The proper extent and limits of public officials acting as agents of the public, and all the ramifications of this issue, are beyond the scope of this paper, except incidentally in discussing the issue of judicial independence. See also, *supra* note 7.

<sup>23</sup>De Tocqueville, in a particularly foreboding and dour state of mind:

It is believed by some that modern society will be always changing its aspect; for myself, I fear that it will ultimately be too invariably fixed in the same institutions, the same prejudices, the same manners, so that mankind will be stopped and circumscribed; that the mind will swing backwards and forwards forever without begetting fresh ideas; that man will waste his strength in bootless and solitary trifling, and, though in continual motion, that humanity will cease to advance.

DE TOCQUEVILLE, *supra*, note 18, Part II, Third Book at 263. Although this passage has fortunately proved to have been unduly pessimistic -- society has certainly advanced in many ways -- the maxim that "the more things change the more they stay the same" is not entirely inapposite here. (Also French, from Alphonse Karr, *Les Guêpes*, Jan. 1849: *Plus ça change, plus c'est la même chose.*)



stature, to its somewhat “hybrid” and not always well-understood nature, to its position outside the judicial branch of government and relatively lower perceived status in the legal hierarchy, and to its history of competing interests, with differing values, trying to mold it to conform to such values. The nature of the subject matter of administrative law no doubt also attracts interest: public policy issues of great concern to many in the government, legal, academic, corporate, and public welfare communities, to name just a few. And of course, as the Honorable Charles W. Burson, now Counsel to the Vice-President, noted in a speech at the 1996 Annual Meeting of the National Association of Administrative Law Judges, “there are huge stakes, in terms of human welfare [and] dollars, [and] there are huge political stakes.”<sup>24</sup>

The issue of performance evaluation of administrative law judges is also illustrative of the unique nature of the administrative adjudicative system, and of some ways in which ALJs and judicial branch judges are perceived differently, above and beyond the obvious fact that the two groups are formally in different branches of government, even if “functionally comparable.”<sup>25</sup> While the issue of evaluating ALJs may also be comparable to the issue of evaluating judges in the judicial branch, in most discussions of evaluation of ALJs significant differences are often presumed, including the degree of control over judges that is implicitly deemed to be acceptable.<sup>26</sup>

---

<sup>24</sup>Tennessee Attorney General Charles W. Burson, Nov. 9, 1996, Address at NAALJ Annual Meeting and Conference, on the subject of “The Future of Administrative Law.” General Burson observed, immediately before making the statement about the huge stakes that are at issue in administrative hearings, that “[i]t is going to be an increasing challenge to maintain your independence against tremendous pressures to conform to the needs of the system -- ‘Look, just get it out, just deny it. You know, if you rule this way, it’s going to disrupt this whole program.’” (A tape-recording of General Burson’s speech is on file with the NAALJ Journal, at Loyola University Chicago School of Law).

<sup>25</sup> See *Butz v. Economou*, 438 U.S. 478, 513 (1978).

<sup>26</sup> Federal ALJs are not evaluated, but are subject to working conditions defined by the agencies that pay them, in a way that judicial branch judges generally are not. See O’Keeffe, *supra*, note 9, at 625; see also Nahum Litt, *Doing It With Mirrors: The Illusion of Independence of Federal Administrative Law Judges*, 36 JUDGES’ JOURNAL 27 (Spring 1997).

However, even in the judicial branch, there have been attempts to oversee the work of judges. Some aspects of such efforts vis-à-vis the federal judiciary have been criticized as constituting a threat to judicial independence by Chief Justice Rehnquist, in his 1995 Year-End Report of the Federal Judiciary (copy on file with the NAALJ Journal, at Loyola University of Chicago School of Law). Addressing the plan of Senator Charles Grassley, Chairman of the

Most significantly, whereas in the judicial branch evaluations of judges are commonly done through bar surveys and/or by specially appointed commissions,<sup>27</sup> in the administrative law arena, although some jurisdictions use survey instruments to obtain feedback from the bar and other interested parties,<sup>28</sup> when performance evaluation of administrative law judges is considered or undertaken, the model is often that of a supervisor evaluating a subordinate employee in a traditional management-by-objectives context, or at least contains aspects of such a process. Such a model would no doubt be considered anathema in the judicial branch, comparable to the impropriety of a superior judge attempting on an ex parte basis to influence another judge with regard to issues in pending or impending cases.<sup>29</sup>

Issues such as these often arise in the ongoing debate about the proper role of ALJs, their independence as decision-makers, and their

---

Senate Judiciary Subcommittee on Administrative Oversight and the Courts, to send questionnaires to all judges asking about the time they devote to judicial and related tasks, the Chief Justice recognized that there was "no doubt that answers to some form of such questions could aid Congress in making decisions about judicial salaries, permitted outside income from teaching, creating new judgeships, and filling existing vacancies," but also observed that "[t]here can also be no doubt that the subject matter of the questions and the detail required for answering them could amount to an unwarranted and ill-considered effort to micro-manage the work of the federal judiciary." *Id.* at 3. The Grassley questionnaire and its dissemination have also been described as being meant as a "shot across the bow" of the federal judiciary, "clearly understood to be such by all parties involved." Francis J. Larkin, *The Variousness, Virulence, and Variety of Threats to Judicial Independence*, 36 JUDGES' JOURNAL 4, 45 (Winter 1997).

The Grassley questionnaires were distributed as planned; the Executive Committee of the Judicial Conference of the United States issued a response in February 1996, noting that the conference was pleased to cooperate fully in the survey and spelling out the "significant cost savings and efficiencies that the judiciary is already achieving"; and the subcommittee issued separate reports on the surveys of the U. S. Courts of Appeal and the U. S. District Courts in May and August, respectively, of 1996. (Copies of the Judicial Conference Response, along with copies of the Subcommittee Reports, obtained from the Administrative Office of the U. S. Courts, on file with the NAALJ Journal, at Loyola University Chicago School of Law.)

<sup>27</sup> See text accompanying note 121.

<sup>28</sup> See, Colorado Division of Administrative Hearings Biennial Report at 25-28, (June 1996). (Copy on file with the NAALJ Journal, at Loyola University of Chicago School of Law.) Colorado's evaluation survey program is described in Edwin L. Felter, Jr., *Administrative Adjudication Total Quality Management: The Only Way to Reduce Costs and Delays Without Sacrificing Due Process*, 15 J. NAALJ 5, 62 (Spring 1995).

<sup>29</sup> See ABA MODEL CODE OF JUDICIAL CONDUCT, Canon 3B(7), printed in full at note 104. See also SHAMAN ET AL., JUDICIAL CONDUCT AND ETHICS, 2d ed., 1995, §5.06, at 159.

accountability to the public. This debate shares many features of the more general conflict today about the proper role of judges, and of the legal system as a whole, in society.

Whether urgent enough to characterize as crisis, the delicate and potentially volatile balance between societal distrust in the legal/judicial system and the legal/judicial system's response thereto, has aspects of crisis: We in the legal/judicial system as a whole, and in the administrative adjudicatory system specifically, may well be in a time of "danger and great difficulty," and at a "decisive moment [or] turning point"<sup>30</sup> with regard to resolution of the pivotal issue of how to protect and maintain judicial independence in a democratic society,<sup>31</sup> and in an imperfect world. As is so often the case with questions of the political ordering of society and its institutions, the resolution of this issue in the context of calls (or, in some jurisdictions, requirements) for traditional evaluation of administrative law judges, may consist not only in what constitutes the best approach, but also in which approach involves the fewest, and the least serious, problems.<sup>32</sup> However, this does not

---

<sup>30</sup>OXFORD DICTIONARY OF CURRENT ENGLISH USAGE (Second Edition 1993) at 200.

<sup>31</sup>See *infra* note 39.

<sup>32</sup>As Utah Supreme Court Justice Christine M. Durham observes in her paper, *Judging in Context: A Discussion Outline*, there really are "no solutions, only difficult choices." *Id.*, at 15 (July 1996). (Copy is on file with the NAALJ Journal at Loyola University Chicago School of Law).

This article centers on evaluation of ALJs by supervisor or chief ALJs, peers, evaluation commissions, and through surveys of the bar and parties who come before ALJs, and not on what probably most ALJs, at least, would agree poses the greatest problems with regard to the evaluation issue: the idea of ALJs being evaluated by agency personnel who may have an interest in the outcomes of cases, which is prohibited for federal ALJs, but which "[a]gencies [have been said to] gaze lustfully at [as] forbidden fruit," and which may exist for some ALJs. See 5 U.S.C. §§4301(2)(D), 4302, 4303 (1982); see also O'Keeffe, *supra*, note 9, at 593-4, 595. The major problem, of course, in this regard is that any interest of such agency personnel in the outcome of cases or classes of cases could obviously affect the neutrality of decision-making by ALJs who are evaluated by such personnel, regardless of whether or not there is any conscious intent to exert any improper influence. Also, persons who are not judges will probably not have the special perspective and awareness of the judicial code of conduct that judges should have, and may thereby overlook important issues of judicial independence.

Those who, like Pierce, believe that agencies not only have the right but the duty to direct ALJ decision-making as a matter of policy, may never be convinced of the extent of the real problems that inhere in agency oversight of ALJ performance, or of the propriety of using a judicial model for ALJ decision-making. See *supra* note 7. However, it is suggested that, for individual litigants, and for the overall system of administrative adjudication, a judicial model is the best last-resort evidentiary decision-making model available. Some better

minimize the importance, or the critical nature, of the questions, the answers, and the manner in which the answers are reached and implemented.

The Chinese word for "crisis" contains two characters: "wei," meaning danger, and "gee," meaning opportunity or chance. This is an apt image and reference point when addressing these primary and essential questions.

The present upheaval -- concerning both the legal/judicial system generally and the administrative adjudicatory system specifically; involving as it does such enormous human and financial stakes and such compelling public policy issues; and occurring in a context of such uncertainty, emotional turmoil,<sup>33</sup> and tendency in this fast-paced world to seek quick solutions to perceived problems -- creates both danger and opportunity: *danger* that, in addressing perceived problems and designing management and evaluation systems for ALJs, concern for popular opinion and/or political considerations will be a primary driving force, and inadequate attention will be paid to fundamental legal principles of due process, judicial independence, and fairness to all parties; and *opportunity* to revisit these basic principles from a fresh

---

alternative means of addressing agency policy concerns and concerns of consistency in decision-making include such straightforward means as rulemaking, along with assuring the continued existence of a high-quality administrative judiciary, through such means as those proposed in this article.

<sup>33</sup>Such emotional turmoil likely includes not only obvious public anger, but also, among other things, less obvious but very simple fear on the part of those under attack that their jobs might be in danger. The popular media has for the past several years been covering stories of the sometimes near-tragic results of "downsizing" in today's economy, and the trend does not appear to be reversing itself yet. As a NEWSWEEK article about the "Dilbert" comic strip phenomenon put it, "Downsizing. Dumb bosses. Double talk. Densification. That's office life in America's favorite comic strip. Too bad reality is even worse." Describing "Dilbert's World," the article continues, "And in the background, burning ever closer are the fires of Competition, triggering the dread drums of Downsizing. 'Knock knock.' says the boss. 'Who's there?' asks the employee. The boss grins: 'Not you anymore!'" NEWSWEEK, Aug. 12, 1996, at 53-55.

Taking just one state as an example, in 1997, Tennessee has abolished 1568 positions, of which 727 were occupied. These numbers do not include positions abolished in 1996 in cutbacks in the department of mental health. The state department of personnel has engaged in ongoing efforts to place all laid-off employees, and had placed 120 as of March 26, 1997, and over 230 as of the end of April, 1997. Interview with Tennessee Employee Relations Director Marianne Batey, Mar. 26, 1997; In Touch, April, 1997 (Monthly publication of the Tennessee Department of Personnel; copy on file with the NAALJ Journal, at Loyola University Chicago School of Law).

perspective, en route to arriving at systems that better preserve both judicial independence and judicial excellence, and also provide for appropriate accountability to the public.

Along the way, it is appropriate to consider the concept of "accountability," in light of both fundamental legal principles and the basic respect all human beings surely owe each other as human beings in a civilized society, at all levels. Modern management theory and research, an area that (to the knowledge of the author) has rarely been visited in discussions of judicial evaluation, offers many helpful insights with regard to the ways in which human beings in organizations interact, "judge" each other, and are held accountable for how well they are perceived to produce.

How best to manage government organizations so as to achieve quality, productivity, efficiency and fairness in serving the public is a question that clearly applies to judicial systems. However, appropriate means of managing judicial systems, including administrative judicial systems, may differ from those that are appropriate for other systems, especially with regard to evaluation of judges. While the stated goals of many traditional performance evaluation programs are worthy -- to provide constructive feedback so that those evaluated can improve their job performance<sup>34</sup> -- the actual effects of ill- or inadequately-considered performance evaluation of judges can not only undermine productivity from a management standpoint, they can also violate the fundamental due process principle of a neutral, independent decision-maker.

---

<sup>34</sup>Scott Adams, creator of the Dilbert cartoons, offers the following perspective: In theory, the performance review process can be thought of as a positive interaction between a 'coach' and an employee, working together to achieve maximum performance. In reality, it's more like finding a dead squirrel in your backyard and realizing that the best solution is to fling it onto your neighbor's roof. Then your obnoxious neighbor takes it off the roof and flings it back, as if he had the right to do that. Ultimately, nobody's happy, least of all the squirrel.

## II. Evaluation of Administrative Law Judges: Fundamental Premises

### A. The American Legal System

The American legal system<sup>35</sup> is based on the idea that, in order to more closely approximate “justice” and effectuate basic fairness, disputes that cannot be resolved informally between parties, or in some alternative agreed-upon manner, should be considered in a somewhat formal context in which all parties have notice of and opportunity to be heard on all issues, and should be resolved based on even-handed, neutral application of the law (case law precedent, statutes, rules), rather than on unfettered discretion, personal feeling and opinion, or political or other inappropriate influence.<sup>36</sup>

This is true because we all have our own views of what is right, fair, and just -- each of us, of course, reasonable in our own eyes, but sometimes strongly disagreeing with others who feel themselves to be just as reasonable. It was no doubt from this truth about human nature

---

<sup>35</sup>The author does not purport to be an expert on the American legal system at the level of some serious scholars on the subject, but, as a lawyer and an administrative law judge, and as a continuing student of the law, has developed some viewpoints on the basic framework and philosophy of the American legal system, and has observed some areas of consensus in this regard.

It is recognized that there are differing philosophies of law. For example, Justice Christine M. Durham lists the following types of legal theories/viewpoints: Critical Feminist Studies, Critical Race Studies, Gay Legal Studies, Legal Process Theory, Post-Structuralism, Pragmatism, Law and Economics, Law and Society, Republicanism, Law and Literature, Post-Modernism, Critical Legal Studies, and Public Choice, in addition to dualism and relativism. Durham, *supra* note 32 at 11. Justice Durham proposes a new pragmatism, or contextualism, as a synthesis of various legal theories; points out that the way we think and know influences how we reach our decisions and affects our judging; and suggests that “[t]o avoid being trapped by our own epistemological paradigms, we should try to view moral and intellectual questions from as many angles as possible.” *Id.* at 2.

The author in this part of this article is attempting to summarize some fundamental legal and human principles on which there appears to be some level of consensus, whatever legal theory or political or other viewpoint one may espouse.

<sup>36</sup>As stated by Ruggero J. Aldisert, Senior United States Circuit Judge for the United States Court of Appeals for the Third Circuit: “We must not establish our conclusions by intense personal desire, keenly felt emotional belief, folklore, superstition, or dogmatic, unquestioning acceptance. Rather, we must state grounds for our conclusion. A conclusion cannot stand on its own direct account, but only on account of something else which stands as ‘witness, evidence, voucher or warrant.’ We have to see an objective connection leading from that which we know to that which we don’t know. . . .” RUGGERO J. ALDISERT, *LOGIC FOR LAWYERS - A GUIDE TO CLEAR LEGAL THINKING*, at 3-3 (1992).

that the idea of following precedent arose -- rather than everything being decided ad hoc, by whoever happened to have the most clout or distinction at the time, the decision-maker was limited by the commonly agreed upon requirement that the decision be based on how other similar situations had been decided in the past. From this, the concept of "equal justice under the law" was developed, through the struggles of many human beings over the space of centuries. And, in the words of Judge Irving R. Kaufman, "Adjudication based on the noble precept 'equal justice under law' requires impartiality, and impartiality demands freedom from political pressure."<sup>37</sup>

To be sure, there are alternative means of dispute resolution, and the development of such alternatives as mediation and arbitration (which often are governed under the legal system) is a welcome development in today's legal community and among the public. However, alternative means of dispute resolution may also include violence,<sup>38</sup> and political means, which unrestrained may lead to tyranny.<sup>39</sup> And,

---

<sup>37</sup>Irving R. Kaufman, *Chilling Judicial Independence*, 88 YALE L. J. 681, 684 (1979).

<sup>38</sup>"People forget that the whole way that English common law started was so that people weren't punching each other out in the middle of the streets. It was a peaceful way to resolve disputes . . . ." Interview with Ellen Hobbs Lyle, newly named Chancellor for Davidson County, Tennessee (who hears appeals from administrative law cases in Tennessee), Gerald Patterson, *Ellen Hobbs Lyle named Chancellor*, NASHVILLE B.J., July 1995, at 3.

<sup>39</sup>See DE TOCQUEVILLE, *supra* note 18, Part I, on the "tyranny of the majority." In Chapter XVI, entitled, "Causes Which Mitigate the Tyranny of the Majority in the United States," the Frenchman counted the legal profession as one of the most important such causes. Observing that "[t]he people in democratic states do not mistrust the members of the legal profession, because it is known that they are interested to serve the popular cause; and the people listen to them without irritation, because they do not attribute to them any sinister designs," *Id.* at 275-276. He further stated, with no discernible trace of irony:

In America there are no nobles or literary men, and the people are apt to mistrust the wealthy; lawyers consequently form the highest political class and the most cultivated portion of society. They have therefore nothing to gain by innovation, which adds a conservative interest to their natural taste for public order. If I were asked where I place the American aristocracy, I should reply without hesitation that it is not among the rich, who are united by no common tie, but that it occupies the judicial bench and the bar.

*Id.* at 278. However, more specifically, on the judiciary, he continued:

[T]he American magistrate perpetually interferes in political affairs. He cannot force the people to make laws, but at least he can oblige them not to disobey their own enactments and not to be inconsistent with themselves. I am aware that a secret tendency to diminish the judicial power exists in the United States; and by most of the constitutions of the several states the government can, upon

however effective alternative means -- both the peaceful and cooperative and the not so peaceful or cooperative -- may be, there must always be a last-resort process for those disputes that parties cannot resolve through alternative means. The American legal system provides a last-resort means of resolving disputes that cannot be resolved any other way. Thus, much in our broader political/governmental system rests on how successful the legal system is in reaching results that stand up over time.

So far, the American legal system appears to be the best model available, in part because it is not a fixed or rigid set of "right answers," but is more a method for finding the answers that are as right as possible for a particular situation, based on what has been done and/or is required to be done in other similar situations. Our legal system is not perfect, and human error is unavoidable even in a good system. However, the legal system is structured to minimize human error through a continuing process of systematic self-correction, which has proven to be more reliable than other, more authoritarian systems, no matter how benevolent the decision-makers.

In our system of justice, it is recognized that deciding cases based on individual, political, or institutional opinions and viewpoints of what

---

the demand of the two houses of the legislature, remove judges from their station. Some other state constitutions make the members of the judiciary elective, and they are even subjected to frequent re-elections. I venture to predict that these innovations will sooner or later be attended with fatal consequences; and that it will be found out at some future period that by thus lessening the independence of the judiciary they have attacked not only the judicial power, but the democratic republic itself.

*Id.* at 279. According to De Tocqueville, "The strength of the courts of law has . . . been the greatest security that can be offered to personal independence." DETOCQUEVILLE, *supra* note 18, Part II, Fourth Book, ch. VII at 325. Reflecting upon these observations today, one is struck not only with the obvious contrast between the trust in lawyers that De Tocqueville perceived and the present distrust (manifested in anger and in "put-down" humor), and by his predictive powers, but also with the possibility that the present turmoil is somehow akin to a popular attempt to overthrow a perceived aristocracy. Whether it will be possible in the face of such popular rebellion to maintain the legal system and the judiciary as truly independent protectors of personal independence and of minorities against the tyranny of the majority, is a question that concerns many. See *supra* note 13. As former Tennessee trial, appellate, and Supreme Court Justice Penny White concluded, "Judicial independence is the backbone of the American democracy[,] essential not only to the preservation of our system of justice, but the preservation of our system of government as well." Penny J. White, *An America Without Justice*, 80 JUDICATURE 174 (Jan.-Feb. 1997).



is right or just in a given situation, without the various procedural protections of due process, is likely to be more subject to biases and prejudices, however unconscious or unacknowledged, than is acting in accordance with due process and other legal principles that have been developed systematically over time and that are equally applicable to all similarly-situated parties. The first fundamental premise of this article is that the administrative legal system, including the setting and enforcing of standards in regard to ALJ performance, should not be excluded or exempted, in whole or in part, from these basic principles that underlie the American legal system.

## B. Independent Decision-Making

### 1. The Nature of Independent Decision-Making

One of the central tenets of our legal system is the due process concept that decision-makers must be independent, in order that they can be neutral and impartial in their decisions. They must avoid, and should be shielded as much as possible from, any influences that might in any way compromise such independence, neutrality, and impartiality -- in order that every person, rich or poor, of whatever standing in the community, can receive equal justice based on the law and not on preconceived notions or improper influences. In our legal system, it is recognized that, despite the best of intentions, a decision-maker who decides a case based in any way on influences and factors outside the facts and law of the case, is more likely to reach an unfair decision than is a decision-maker who makes a serious effort and commitment to limit his or her inquiry to the facts and law of the case, in accordance with ethical and other requirements imposed by law, independent of outside influences, and even of the judge's own personal opinion, when it conflicts with the law.

Indeed, a judge, Justice Benjamin Cardozo observed, "is not to innovate at pleasure. He is not a knight-errant roaming at will in pursuit of his own ideal of beauty or of goodness."<sup>40</sup> As former Chief Justice Roger J. Traynor of California cautioned: "[O]ne entrusted with

---

<sup>40</sup>BENJAMIN N. CARDOZO, THE NATURE OF THE JUDICIAL PROCESS, 141 (1921).

decision, traditionally above base prejudices, must also rise above the vanity of stubborn preconceptions, sometimes euphemistically called the courage of one's convictions. [A judge] must severely discount his own predilections, of however high grade he regards them, which is to say he must bring to his intellectual labors a cleansing doubt of his omniscience, indeed even of his perception."<sup>41</sup> More recently, Justice Breyer has summarized "[t]he question of judicial independence [as] revolv[ing] around the theme of how to assure that judges decide according to law, rather than according to their whims or to the will of the political branches of government."<sup>42</sup>

Unfair decisions may also occur "innocently," as a result of a judge's lack of experience similar to that of litigants, or as a result of unconscious biases and reactions,<sup>43</sup> but the results for litigants may be just as serious as if they were caused by blatant prejudice or clearly coercive influence. Thus, it is important for judges to become as aware as possible of our own biases, prejudices, propensities, and ways of knowing and thinking, in order to avoid allowing these to improperly influence outcomes in the cases that come before us.

As Justice Cardozo recognized, "We may try to see things as objectively as we please. None the less, we can never see them with any eyes except our own."<sup>44</sup> We do have the law to guide us, but more is required, in terms of our own self-awareness:

Much of the law is designed to avoid the necessity for the judge to reach what Holmes called his 'can't helps,' his ultimate convictions or values. The force of precedent, the close

---

<sup>41</sup> Roger J. Traynor, *Reasoning in a Circle of Law*, 56 VA. L. REV. 739, 750 (1970), quoted in ALDISERT, *supra*, note 36, at 2-14.

<sup>42</sup>Breyer, *Judicial Independence in the United States*, 40 ST. LOUIS U. L. J. 989 (1996), cited in Larkin, *supra*, note 26, at 4.

<sup>43</sup>See Charles R. Lawrence III, *The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism*, 39 STAN. L. REV. 317 (1987), for a discussion of unconscious discrimination in the context of the requirement of proving a racially discriminatory purpose in constitutional challenges to facially neutral laws. Referring to "the multitude of parochial self-interests the unconscious seeks to disguise and shield," Professor Lawrence, of Stanford Law School, argues that "judicial exploration of the cultural meaning of governmental actions with racially discriminatory impact is the best way to discover the unconscious racism of governmental actors." *Id.* at 387-388.

<sup>44</sup> CARDOZO, *supra* note 40, at 13. See also Durham, *supra* note 35.

applicability of statute law, the separation of powers, legal preemptions, statutes of limitations, rules of pleading and evidence, and above all the pragmatic assessments of fact that point to one result whichever ultimate values be assumed, all enable the judge in most cases to stop short of a resort to his personal standards. When these prove unavailing . . . the judge necessarily resorts to his own scheme of values. It may therefore be said that the most important thing about a judge is his philosophy; and if it be dangerous for him to have one, it is at all events less dangerous than the self-deception of having none.<sup>45</sup>

If we are to be truly independent decision-makers, we must acknowledge and address the potential effects of such underlying values and forces. Professor Richard Pierce quotes Judge Jerome Frank on this subject:

Interests, points of view, preferences, are the essence of living. . . . The conscientious judge will, as far as possible, make himself aware of his biases . . . and, by that very self-knowledge, nullify their effect. Much harm is done by the myth that, merely by [becoming] a judge, a man ceases to be human and strips himself of all predilections. . . . The concealment of the human element in the judicial process allows that element to operate in an exaggerated manner; the sunlight of awareness has an antiseptic effect on prejudices.<sup>46</sup>

Judge Frank and Professor Pierce raise a valuable point: In some ways, perhaps *because of* the strong need to project fairness and avoid even the appearance of impropriety, judges may actually be *more* subject to subconsciously concealing and thereby actually exaggerating any underlying human preferences, values, and biases. Also, because judges are regularly called upon to resolve the disputes of others, we

---

<sup>45</sup>Paul Freund, *Social Justice and the Law*, SOCIAL JUSTICE 93, 110 (R. Brandt ed. 1962), as quoted in ALDISERT, *supra* note 36, at 2-12.

<sup>46</sup>Pierce, *supra* note 7, at 519, quoting from *In re J. P. Linahan, Inc.*, 138 F.2d 650, 651-53 (2d Cir. 1943).

may develop a propensity to believe that we are called upon to use our own personal wisdom to help others resolve their disputes, and indeed to believe in the superiority of such personal wisdom. We may neglect to maintain awareness of personal biases, values, and preferences, and we may even become self-satisfied and appear to be arrogant. This quite obviously can contribute to the sort of public outrage against judges that we see today.

Of course, even the brightest sunlight may not be completely effective in disclosing and eradicating prejudices and biases. As Redish and Marshall note in their article, *Adjudicatory Independence and the Values of Procedural Due Process*, “[i]n many instances the pressures on an adjudicator may be so subtle that not even she is aware that her decision has been shaped by improper influences.”<sup>47</sup> It is suggested that any judges who claim to be immune to improper influences, or to have erased all prejudice, bias and predilections from their psyches, may well be fooling themselves, and fall into Freund’s most “dangerous” category. Maintaining awareness of the possibility that any of us might be, even unconsciously, affected by such forces -- both external and internal -- would seem to be a key to being a good judge who practices independent decision-making in the best sense of the words.

## 2. The Responsibility of Independent Decision-Making

The second fundamental premise of this article is that independent decision-making should indeed be viewed less as a power than as an indispensable responsibility of all judges, at all levels, including ALJs. At times, in discussions of independent decision-making, some participants, both pro and con on the issue, seem to take for granted that the issue has to do with the level of personal power or authority an individual in a judicial or quasi-judicial role enjoys.

This approach, while it recognizes some basic truths about human temptation, is flawed, because the pertinent issue is not really how much power or authority an individual in a judicial role should have for himself or herself, but how much practical independence and authority on the part of a decision-maker is required for effectively fair and

---

<sup>47</sup> Martin H. Redish & Lawrence C. Marshall, *Adjudicatory Independence and the Values of Procedural Due Process*, 95 YALE L. J. 455, 494 (1986).

neutral decision-making. University of Pennsylvania Professor Stephen B. Burbank put it thus: "As obvious as this may seem, there is a risk, to which even federal judges may fall prey from time to time, of reifying the concept of judicial independence and of treating it as a disembodied goal rather than the means to higher ends."<sup>48</sup> ALJs as well will do well to guard against falling into this logical fallacy.

True independent decision-making is based on the law and facts of cases,<sup>49</sup> as independent as possible of any internal biases, external pressures, or information not equally accessible to all parties. True independent decision-making has less to do with personal power and the trappings thereof than it does with fulfilling the values of due process. Indeed, true independent decision-making is not easy, and the author would argue that it should not be taken for granted, or sought, as a prerogative to be "enjoyed."

The difficult and even onerous nature of true independent decision-making is perhaps best illustrated by reference to the sometimes extreme negative consequences suffered by some federal judges in the civil rights era of the 1960's.<sup>50</sup> It may reasonably be supposed that these judges experienced their independence less as something they enjoyed than as something they felt a deep commitment to fulfill, despite the shunning and more serious consequences they suffered as a result.

Independent decision-making involves more subtle personal sacrifice as well. As illustrated in the preceding section, it demands paying rigorous attention to deciding cases based not on a judge's personal beliefs and convictions on what is right or appropriate (which might otherwise seem natural and feel "right") but on as neutral as possible interpretation and application of the law, after hearing the evidence and arguments of all parties, whether or not the judge agrees with the law. It requires resisting political, personal, financial and other forms of pressure at a level not expected of non-judges. It involves

---

<sup>48</sup>Burbank, *supra* note 20, at 117.

<sup>49</sup>The principle of the exclusiveness of the record has been called "the basic principle governing administrative hearings and decisions." Bernard Schwartz, *Administrative Law Cases During 1995*, 48 ADMIN. L. REV. 399, 405 (1996). See *Goldberg v. Kelly*, 397 U.S. 254, (1970): "Finally, the decision maker's conclusion . . . must rest solely on the legal rules and evidence adduced at the hearing." *Id.* at 271.

<sup>50</sup>See Kaufman, *supra* note 37, at 689-690.

avoiding much of the kind of informal discussion with non-judges that others naturally engage in about their work.

### 3. The Importance of Independent Decision-Making

Redish and Marshall posit that an independent adjudicator is the necessary "floor" or core element of due process, without which the values of due process cannot be realized, in any legal proceeding -- whether in court or in an administrative context.<sup>51</sup> They contend that "the one procedural protection that is clearly necessary for the fulfillment of all the goals of due process is the participation of a truly independent adjudicator." Stating that "the participation of an independent adjudicator is such an essential safeguard" that the value of due process "cannot be protected" without it, they find judicial independence to be so basic that it may be the "only" essential safeguard of due process.<sup>52</sup> Other safeguards such as the right to notice, hearing, counsel, etc. "are of no real value . . . if the decisionmaker bases his findings on factors other than his assessment of the evidence before him."<sup>53</sup>

The concept of an independent decision-maker is therefore a fundamental one in the American legal system, including the administrative legal system. And although ALJs may not have the same level of authority, power, or practical independence as judicial branch judges, they must adhere to the same standard of impartial decision-making.<sup>54</sup> Moreover, it has been held that, when procedural safeguards available in court proceedings are not present in administrative proceedings, the due process requirement of an impartial decisionmaker is to be applied even more strictly than in court proceedings.<sup>55</sup>

---

<sup>51</sup>Redish, *supra* note 47, at 456-457.

<sup>52</sup>*Id.* at 475

<sup>53</sup>*Id.* at 476.

<sup>54</sup>*See* KOCH, *supra* note 7, §6.7[2] at 218; Barry v. Bowen, 825 F.2d 1324, 1330 (9th Cir. 1987), *appeal after remand* 884 F.2d 442 (9th Cir. 1989); Kendrick v. Sullivan, 784 F. Supp. 94, 102 (S.D.N.Y. 1992).

<sup>55</sup>Ventura v. Shalala, 55 F.3d 900, 902 (3d Cir. 1995). Some of the types of less formal on-the-record administrative proceedings that exist at the federal level are described in William Funk, *Close Enough for Government Work?--Using Informal Procedures for Imposing Administrative Penalties*, 24 SETON HALL L. REV. 1 (1993). Funk addresses the agency

To guide us in meeting this responsibility, codes of conduct for ALJs,<sup>56</sup> as for judicial branch judges, in many jurisdictions contain numerous requirements conceived to protect and ensure our independence. These range from limitations on financial and political activities, to rules on conflicts of interest and disqualification, to the ex parte prohibitions, which exist to ensure that decisions are made based solely on the record and the law and not on any outside influences or information, to the broad imperative to avoid even the appearance of impropriety.

Any management and/or evaluation program for ALJs must be very carefully crafted and implemented in light of these requirements, which can assist us in becoming the judges we should be. That we need help and guidance in this regard can hardly be disputed. For human beings are not computers able to exclude completely from consideration specified factors, but (notwithstanding some theoretical few bastions of strength among us) are at bottom fallible creatures, affected by pressures tremendous and small, especially when they are perceived -- consciously or unconsciously -- as constituting any sort of threat to life, limb, or happiness. And as Professor Rosenblum has observed, in the context of performance evaluation, "People's livelihoods are enhanced or decimated depending upon whether supervisors check 'good' or 'poor' on a standardized evaluation form."<sup>57</sup>

---

motivations for desiring such informal procedures -- both legitimate cost, time, and case management considerations, as well as the more covert control motive to win cases more easily, *Id.* at 67-68; notes the erosion of the intended broad applicability of the federal APA and uniformity of procedures under the Act; and suggests that if informal procedures are to be used, the APA should be amended in order to provide more appropriate uniformity in such proceedings. *Id.* at 66-69. Professor Funk also observes that "the lack of assured neutrality of the Presiding Officer [in such less formal proceedings] must be addressed," and suggests that those who hear such cases, whether they be ALJs or "presiding officers," should be better insulated "from both the reality and appearance of enforcement oversight and influence," through such means as "making the job fulltime and not having the Presiding Officer rated by persons in the enforcement office . . . ." *Id.* at 61.

<sup>56</sup> See note 109, *infra*.

<sup>57</sup> Rosenblum, *supra* note 1, at 1. Although the focus of this article is not on how aspects of evaluation systems in the judicial branch of government may potentially compromise judicial independence, it is interesting to note that Rosenblum goes on, immediately after the language quoted in the text, to state that "Bar associations influence the makeup of our judiciary by releasing results of evaluative bar polls just prior to elections." *Id.*

## C. Management Principles

### 1. Finding an Appropriate Balance: Maintaining Independence and Standards of Performance

Professor Rosenblum's observation illustrates the dangers that are inherent in evaluating administrative law judges. If livelihoods can be "decimated" by negative evaluations, simple principles of human nature would suggest that people being evaluated on their work performance will tend to try to do what is necessary to obtain more rather than less positive marks on such evaluations. As noted by Redish and Marshall, quoting Alexander Hamilton in the Federalist Papers, "[i]n the general course of human nature, *a power over a man's subsistence amounts to a power over his will.*"<sup>58</sup>

Such power may not always exert itself on a conscious level. If judges are subject to internal influences and motivations of which they may not be very aware, and to "forces which they do not recognize" tugging at them and affecting their decisions,<sup>59</sup> it is likely that they, like others who are evaluated, may well on some level of consciousness try to do what is necessary to obtain more rather than less positive evaluations, even when they would otherwise act differently, all other things being equal. This may be appropriate, when the changes in behavior are in accordance with the law and what is generally agreed to constitute good judging. However, other changes in behavior may occur as well, which may not be so justifiable.

As Hamilton suggested, inappropriate changes in behavior may be most likely when the results of evaluations can have any effect on judges' livelihoods. The use of evaluations in judicial systems in which judges are elected can obviously affect re-election chances. For administrative law judges in many jurisdictions, evaluations are directly tied to "merit pay" increases or "pay for performance." In addition, evaluations may affect order of layoff in a reduction in force.<sup>60</sup> Finally,

---

<sup>58</sup>Redish, *supra*, note 47 at 481, quoting THE FEDERALIST No. 79, at 583 (Alexander Hamilton) (E. Earle ed. 1937) (emphasis in original).

<sup>59</sup>CARDOZO, *supra* note 40, at 12-13. See also text accompanying notes 40-47.

<sup>60</sup>See, e.g., TENNESSEE CODE ANN. §8-30-320(a)(2), which allows performance ratings to be considered in determining order of layoffs "when the seniority calculations produce an order of layoff difference of less than one (1) year."



to the degree evaluations are public, they can affect a judge's reputation, and thereby his or her future career. Given such realities, it is possible that acting in such a way to improve one's chances of receiving good evaluations may well at times include acting in a less rather than more fair, efficient and otherwise legally appropriate manner.

In a perfect world, there might be a perfect evaluation system in which perfect human beings would evaluate other human beings, including administrative law judges, purely on the basis, for example, of how fairly and efficiently -- according to perfectly-defined standards of fairness and efficiency -- those evaluated have performed their duties. And those evaluated would have no motivation other than to become more fair and more efficient. However, we do not live in a perfect world, and we must consider such realities as the limitations of human intelligence, "human nature," deep-seated biases and prejudices -- both obvious and subtle, blatant and unacknowledged -- and the tendency to respond to perceived threats to one's well-being, even against one's better judgment.

The question becomes: How should any standards setting and enforcement in an adjudicatory system be structured, from a legal and management standpoint, so as to minimize inappropriate responses, maximize fairness and neutrality in decision-making, and best assure that appropriate standards of competence and conduct are maintained?

There is, of course, an inherent tension between the concept of independent decision-making and maintaining standards of competence and conduct, if anyone other than the individual decision-maker is to participate in the latter. With complete independence, there would be no outside participation in setting or enforcing any standards of conduct or competence and, except for the corrective and possibly restraining effects of the appellate process, an incompetent judge who engaged in improper conduct would have free rein to inflict harm on parties, and on the system. On the other hand, if inappropriate standards are set, or enforcement is biased, too heavy-handed, or otherwise compromised by inappropriate factors or procedures, independent decision-making and fairness to parties and the system may suffer.

---

An appropriate balance must be found: if fair, impartial, well-reasoned, and timely decisions are the desired outcome, an adjudicative system should be organized to foster as much practical independence as possible on the part of the decision-makers, while at the same time ensuring that appropriate standards are set and enforced, in an appropriate manner. To preserve the integrity of such a system, the job of judge should be structured so as to make fulfilling the responsibility of making fair, impartial, well-reasoned, and timely decisions as easy as possible, and not more difficult. If there is to be any evaluation of judges, it should be done in such a way to facilitate rather than hinder judges in fulfilling the responsibility of independent decision-making. To this end, management principles that promote the values of the American legal system and independent decision-making should guide the management of administrative adjudicative systems, including any evaluation, broadly defined as any standards-setting and enforcement, in regard to ALJs.

## 2. Modern Management Thinking

The fact that most ALJs are hired rather than appointed for a term or elected, and are considered to be “employees” in a government bureaucracy, has led to the use of management-by-objectives approaches with regard to many ALJs. However, according to several management experts who have raised some provocative and compelling questions about some previously little-questioned assumptions, traditional management by objectives and traditional periodic performance evaluation systems, although deeply ingrained in American management philosophy, may not serve us well, especially with regard to work that is difficult to quantify objectively and that requires independent thought and decision-making. According to the theories of these experts, traditional performance evaluation systems are inherently subjective exercises and can actually diminish effective performance in any work setting. Professor Canice Prendergast of the University of Chicago Graduate School of Business has illustrated how such evaluation systems and other ranking mechanisms tend to create

“yes” men (and women),<sup>61</sup> and are subject to discrimination and bias that is difficult to monitor.<sup>62</sup>

Prendergast observes that in real-world employment relations, “[s]upervisors use subjective information to evaluate subordinates’ performance and to allocate rewards, but supervisors are not themselves the residual claimants of subordinates’ output. This leaves room for supervisors’ preferences, and biases, to affect rewards by manipulating the appraisal system.”<sup>63</sup> Prendergast notes a 1985 survey, which established that supervisors give higher ratings to subordinates of their own race -- the race of both the raters and the ratees affected evaluations.<sup>64</sup> It has also been shown that, in judicial evaluation programs, women judges endure consistently stronger criticism than their male colleagues, especially in particularly subjective categories such as demeanor. Those women judges who score high in the areas of legal knowledge, promptness, and case management are subjected to “condescending barbs” and “often receive low marks for strong and decisive action, behavior that garners praise for their male colleagues.”<sup>65</sup>

Whether such bias is intentional or unconscious, the effects remain the same, and obviously compromise the integrity of evaluation systems. And sometimes, clear favoritism may be at work. For example, Prendergast cites research illustrating the political aspects of performance appraisal, giving an example in which Navy supervisors were found to evaluate favored subordinates so as to maximize the likelihood of their promotion.<sup>66</sup>

Bias can cause inefficiencies on a number of dimensions: Prendergast notes that “[e]mployees who feel discriminated against may quit, with resulting turnover costs and lost human capital for the organization.”<sup>67</sup> Discouraged effort among those who conclude that

---

<sup>61</sup>Canice Prendergast, *A Theory of “Yes Men,”* 83 AM. ECON. REV. 757 (1993).

<sup>62</sup>Canice Prendergast & Robert Topel, *Discretion and Bias in Performance Evaluation*, 37 EUR. ECON. REV. 355 (1993).

<sup>63</sup>*Id.* at 356.

<sup>64</sup>*Id.* at 358.

<sup>65</sup>*Unfinished Business: Overcoming the Sisyphus Factor, A Report on the Status of Women in the Legal Profession*, ABA Comm. on Women in the Profession, at 17 (1995).

<sup>66</sup>Prendergast, *supra* note 62 at 358.

<sup>67</sup>*Id.* at 359.

they are out of the running for rewards can offset any encouragement of the front-runners.<sup>68</sup> Bias “makes it difficult to distinguish genuinely good performance from favoritism.”<sup>69</sup> Incentives for effort are actually reduced by such bias, according to Prendergast, both in the favored and the unfavored, by lowering morale and thereby leading to lower effort and output overall. It also leads to interpersonal “influence activities” and subtle forms of bargaining, or “rent-seeking,” with superiors.<sup>70</sup>

Monitoring bias is difficult, according to Prendergast, because supervisors have an information advantage. Also, supervisors have control over many decisions that affect subordinates’ productivity -- through assignments, provision of training, informing them of opportunities, etc. Monitoring can raise the possibility “that supervisors will ‘sabotage’ the performance of workers in order to justify their biased ratings.” And subjects of evaluation often fear reprisals from supervisors if they report them for unfair treatment. Empirical work cited by Prendergast shows that management is reluctant to reverse decisions made by lower-level supervisors, not wanting the supervisors to “lose face,” or to suffer “the negative externality caused by decision reversal on future effort incentives” by workers who lose trust in the reversed supervisor. For all these reasons, according to Prendergast, the exercise of favoritism is extremely difficult to constrain.<sup>71</sup>

Prendergast concludes that subjectivity is central to performance appraisal in most organizations, and that bias in evaluations can give rise to inefficiencies, not only by negatively affecting the performance of workers, but also by making it difficult to determine the true talents of workers.<sup>72</sup>

An additional aspect of the problem of bias and favoritism is that supervisors have been shown to distort their evaluations more when their decisions substantially affect subordinates’ welfare. Also, favoritism generates value for supervisors, and leads naturally to a demand for “power” by those in authority, according to Prendergast and

---

<sup>68</sup>*Id.* at 362.

<sup>69</sup>*Id.* at 359.

<sup>70</sup>*Id.*

<sup>71</sup>*Id.* at 360-361.

<sup>72</sup>*Id.* at 364.

Topel in a more recent paper on favoritism in organizations.<sup>73</sup> Noting that the economics literature in the past has generally ignored issues such as personal preferences toward employees and the demand for power by management,<sup>74</sup> Prendergast and Topel provide valuable analysis in understanding the subjective nature of performance evaluations.

In his "theory of yes men," Prendergast points out that "when the point of reference for adequate performance is the manager's opinion, an endogenous desire arises for the worker to conform to the opinion of the manager. This arises because the only way to induce the worker to exert effort is by comparing the findings of the manager with those of the worker. Hence subjective performance evaluation gives rise to the existence of 'yes men,' who attempt to second-guess the opinions of their monitor and mimic them."<sup>75</sup>

This goes further, and leads to a desire for conformity, with workers reporting and acting not just on the basis of their own observations, but also on what they believe the opinions of others to be. "Hence interaction stifles 'creativity,' as all workers converge inefficiently to a similar conclusion." Prendergast compares his conclusions to other

---

<sup>73</sup>Canice Prendergast & Robert H. Topel, *Favoritism in Organizations*, 104 J. POL. ECON. 958, 959 (1996).

<sup>74</sup>*Id.* at 975. Prendergast and Topel observe that "bribery and lobbying [addressed by other studies] are unnecessary to illustrate the [issue]; all that is necessary is that supervisors have likes and dislikes toward their subordinates." Further, they posit two additional inefficiencies, notwithstanding the benefits -- to supervisors -- of greater authority: increased risk on workers and more inefficient task assignments. *Id.* at 976.

Another study by Prendergast (with Lars Stole) has to do with the manner in which decision-makers approach new information when decisions reflect on people's ability to learn. Although their study was done in the context of a manager making investment decisions over time, some of their findings may have relevance in an evaluation context. Prendergast and Stole identify two possible distortions in responses to new information, which arise from a wish to acquire a reputation as quick learners, and which coincide with the theories of social psychologists: first, exaggerated over-responses at the early stages of receiving new information, which corresponds to the "base rate fallacy or overconfidence effect" of social psychology; and later, conservatism, corresponding to the concepts of "cognitive dissonance reduction" and the "sunk cost fallacy," in which individuals commit more resources to a losing cause so as to justify or rationalize their previous behavior. If such responses come into play in evaluation situations, in which it may be said that there may be some incentive for acquiring a reputation for learning, then the actual value of evaluation programs may obviously be less than predicted. Canice Prendergast & Lars Stole, *Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning*, 104 J. POL. ECON. 1105, 1106, 1125 (1996).

<sup>75</sup>Prendergast, *supra* note 61, at 758.

literature in the field of management, on principal-agency relationships, and on “herding” and conformity.<sup>76</sup> He concludes that “yes” men are likely to be concentrated among the less able workers, among workers with less able managers, in organizations with much interaction between management and workers, and in organizations with “high-powered incentives.”<sup>77</sup> However, he notes “one key point that must be stressed:”

I have not assumed that workers wish to mimic their supervisors per se. Instead, this arises endogenously as the only way to induce the worker to exert effort so that the existence of yes men is a natural implication of subjective performance evaluation.<sup>78</sup>

Peter Block, a management expert and consultant, has also observed the “fundamentally subjective” nature of performance evaluations, suggesting that, “[g]iven the subjective nature of evaluations, we are as likely to be rating and paying people for compliance as we are for performance.”<sup>79</sup> More bluntly, Block terms appraisal “a process of coercion.” Noting that “[w]e also call it a reward system. Yet if it is a reward system, it is a punishment system too.”<sup>80</sup> He asserts that “[i]f the intent of the appraisal is learning, it is not going to happen when the context of the dialogue is evaluation and judgment.” Acknowledging that the appraisal process is a “tough one to let go,” Block advocates self-evaluation, evaluation by “customers” rather than by bosses, and less formal means of feedback.<sup>81</sup>

One of the earliest thinkers in the field of management who raised questions about performance evaluations was W. Edwards Deming, who in the period after World War II worked with Japanese corporations, teaching them how to be productive and maintain quality. Deming has asserted that management by objectives, the prevailing

---

<sup>76</sup>*Id.* at 759; (citations omitted).

<sup>77</sup>*Id.* at 769-770.

<sup>78</sup>*Id.* at 763.

<sup>79</sup>PETER BLOCK, *STEWARDSHIP -- CHOOSING SERVICE OVER SELF-INTEREST*, 172 (1993).

<sup>80</sup>*Id.* at 199.

<sup>81</sup>*Id.* at 152-154.

theory of which performance evaluation is a part, and ranking employees, is “devastating” to true productivity, and should more accurately be called “management by fear.” According to Deming, such a system nourishes short term performance, at the expense of effective long term planning. It also “builds fear, demolishes teamwork, nourishes rivalry and politics [and] leaves people bitter, crushed, bruised, battered, desolate, despondent, dejected, feeling inferior, some even depressed, unfit for work for weeks after receipt of rating, unable to comprehend why they are considered to be inferior.” Further, it is “unfair, in that it ascribes to the people in a group differences that may be caused totally by the system that they work in.”<sup>82</sup>

Observing that people who are evaluated “are afraid to ask questions that might indicate any possible doubt about the boss’s ideas and decisions, or about his logic,” Deming suggested that “[t]he game becomes one of politics. Keep on the good side of the boss. Anyone that presents another point of view or asks questions runs the risk of being called disloyal, not a team player, trying to push himself ahead.” In the same vein as Prendergast, Deming noted the pressure to “[b]e a yes man,” and concluded that fair rating is impossible.<sup>83</sup>

Deming also found the setting of numerical goals and quotas to be non-productive, because they can lead to distortion and faking. One example he gave was that of a federal mediator who was rated on the number of meetings he conducted during the year; the mediator had improved his rating by stretching out to three meetings negotiations between parties when he could have settled all the problems in one meeting.<sup>84</sup> Other examples can easily be posited: ALJs evaluated on numbers of cases completed taking less time than appropriate to consider evidence and analyze it is light of relevant law, being more likely to hold parties in default and dismiss (and thereby complete) cases when it may not be clear the parties received notice of a hearing,<sup>85</sup>

---

<sup>82</sup>W. EDWARDS DEMING, OUT OF THE CRISIS, 102.( 1982)

<sup>83</sup>*Id.* at 108-109.

<sup>84</sup>*Id.* at 106.

<sup>85</sup>This example was given by one of the presenters at a seminar presented by the ABA National Conference of Administrative Law Judges held October 12, 1996, in Chicago, on the subject, “Surviving the Pressure Cooker: A Workshop to Help Judges Maintain the Balance Between Agency Objectives and Judicial Independence.”

etc.

Deming advocated the abolishment of all ranking and other barriers to taking pride in one's work; the encouragement of communication and continual learning; the study and management of systems, to improve processes; and analysis of the distinction between common causes of variation and special causes, in order to better understand the kinds of action to take. He espoused cooperation rather than competition among workers, improving systems rather than targeting perceived faults in workers, and avoiding reliance on extrinsic motivation in favor of intrinsic motivation, self-esteem, and dignity. He believed that such a transformation of leadership concepts would result in higher quality and productivity in the workplace.<sup>86</sup>

### 3. The Relevance of Modern Management Thinking to Considerations of Accountability

Some of the characterizations of performance evaluations summarized above may sound extreme to those who believe strongly in performance evaluation as a means of maintaining accountability with the public. However, accountability to the public is a somewhat slippery concept in this day and age, in contexts in which issues are not easily definable in short sound bites. Care must be taken to be sure that what is desired is not merely the appearance of accountability, to calm a restless and angry (and voting) public.

True accountability in the context of a judicial system, including an administrative judicial system, should include above all assuring the public that proceedings are fair. It should also include assuring the public that proceedings are conducted in a timely manner, and efficiently, but not at the expense of fairness or depriving parties of their right to be heard. It should include assuring the public that those who serve as judges are qualified, and that if there are problems with a judge's performance that arise to a level of significance generally agreed, in an appropriate context, to constitute an appropriate basis for removal, those problems will be corrected or the judge will be removed. Finally, in order to achieve these other goals of accountability, it should

---

<sup>86</sup>W. Edwards Deming, *Leadership for Quality*, 11 EXEC. EXCELL. 3-5 (1994) (published after Deming's death in December 1993).



include public education on how various parts of the legal system work.<sup>87</sup> Only with such education will the public likely come to understand and support the concept of judicial independence.

Judicial performance evaluation systems have been and will continue to be adopted, to address public concern over judicial accountability. However, to ignore the knowledge of experts in the field of management on the subject of performance evaluation, a process that was developed in the field of management, would seem to be foolhardy. Moreover, surveys have shown that most persons who participate in evaluations, both supervisors and employees, "rate the process a resounding failure," according to a November 1996 Wall Street Journal article.<sup>88</sup> Citing studies concluding that up to 90 percent or more of performance evaluation systems used by business are unsuccessful, the article listed as reasons for this that the process is often "just a ritual," that systems are poorly designed, that supervisors dislike confrontation and employees often react negatively (noting one incident in which a worker killed four of his bosses after receiving a poor review), and that the process is simply ineffective at improving performance. Problems in multi-rater review systems -- which include evaluation by peers, subordinates and customers in addition to supervisors -- were also identified: Such systems "often become a popularity contest" and "produce rabid politics and widespread dissatisfaction."<sup>89</sup>

---

<sup>87</sup> See *supra*, note 17.

<sup>88</sup> Timothy D. Schellhardt, *Annual Agony*, WALL ST. J. November 19, 1996, at A1.

<sup>89</sup> *Id.* at A10. Alternatives to traditional evaluation that are discussed in the article include more frequent reviews, and giving up forms altogether in favor of teaching everyone how to give and receive good feedback. *Id.* Another alternative might be to take performance evaluation to its logical end, which might be something like one extremely rigorous form practiced by a major oil company. The company is organized into a hierarchy of various levels of supervisors, with each supervisor responsible for annually evaluating approximately 10 persons who work under him or her. Everyone at all levels is subject to this process. Each supervisor must rate his or her supervisees in order of rank, according to how well each supervisee is seen to be producing. Then, groups of approximately 10 supervisors at each level meet, and merge their separately-ranked lists into combined ranked lists of some 100 employees each. This is done through a process of comparison of the supervisors' experiences with the various supervisees; disagreement and open discussion is encouraged, in order to reach accurate determinations, and efforts are made to take relevant variables into account so as to reach fair decisions on ranking. If one supervisor has little success with a particular supervisee, that employee will be placed under another supervisor for a period of time to determine whether he or she might be more productive under someone else. Persons who end

Given these dangers, and given that a significant problem identified by the management experts is that performance evaluation systems compromise independent thinking, the judicial community would do well to attend to the information these experts have to offer, if a system with true accountability and integrity is the goal. While judges are subject to numerous requirements intended to preserve independence, pressures from the public are becoming more insistent, with the potential for having real impacts in judges' lives, which can compromise judicial independence despite conscious efforts to avoid such compromise.<sup>90</sup> It cannot be assumed that, by some magical powers, judges above all other human beings will resist all pressures that might affect their jobs and their livelihoods. Therefore, to adopt well-intended programs that nonetheless may have the potential to add to these pressures would not appear to be wise.

The question, again, is: how should any standards setting and enforcement in an adjudicatory system be structured, from a legal and management standpoint, so as to minimize inappropriate responses, maximize fairness and neutrality in decision-making, and best assure that appropriate standards of competence and conduct are maintained? And further, are there any unique qualities about administrative law judge offices that would suggest approaching this issue any differently in such offices? Before addressing these questions, it is appropriate to examine the traditional model for evaluation of many ALJs in light of the three premises discussed above, and in light of what we can learn from the management experts. This examination will start with a consideration of the intended purposes of traditional performance evaluation, and then move to an analysis of some of the unintended consequences of implementing a traditional model with ALJs.

---

up in the bottom five to ten percent of the merged lists more than two times are terminated from employment. In this manner, those determined to be the least able and productive, by a number of supervisors over time, are screened out, and those perceived to be the most able and productive are promoted. March 30, 1997, personal interview with current employee of major oil company.

<sup>90</sup>See *supra*, text accompanying notes 40-47, 57.

### III. Intended Purposes and Unintended Consequences of Traditional ALJ Evaluation Systems

#### A. Purposes of Performance Evaluation

The purposes of performance evaluation systems are generally good and well-intentioned, and include such things as productivity, efficiency, and quality and consistency of work product.<sup>91</sup> Without question, the image of an incompetent, lazy, biased, and/or unethical government worker or official at any level, including judges and ALJs, is an abhorrent one, and indeed is probably more abhorrent to committed public servants than it is to the public at large. Such persons make the rest of us look bad, by association. Quality, productivity, efficiency, and fairness are not simply desirable outcomes; they are necessary, and they apply with equal if not greater consequence in judicial systems, including administrative judicial systems.

Performance evaluation is seen and promoted as an effective means of achieving these worthy and important ends. By giving those who are evaluated notice of expected performance and feedback on how well they are performing, performance evaluation is designed to produce improved performance in the majority of those evaluated, and to identify individuals who either need special assistance to meet appropriate performance standards, or must be expelled from the work force in order to maintain its quality and integrity, according to supporters of evaluation. Whether and how well traditional performance evaluation programs achieve these ends is open to question, however, from a management standpoint, as the previous section of this article illustrates. The next section discusses some of the ways in which there may be unintended negative consequences in applying a traditional performance evaluation model in an administrative judicial context, from a combined legal and management standpoint.

---

<sup>91</sup>See *supra* notes 5-8, and accompanying text.

## B. Unintended Consequences of Traditional Performance Evaluation

No matter how well-intentioned their purposes, the unintended negative consequences of traditional evaluation systems -- in which ALJs are ranked and in which pay and job security may be affected by evaluation scores -- can sometimes be serious, and can rise to a level that their negative effects outweigh any good that may be achieved through their use. Such negative effects can include compromised independence in decision-making, lowered morale and quality of work, and decreased efficiency and productivity.

### 1. Compromised Independence

The primary, central, and critical negative consequence of evaluation of ALJs is, quite simply, the very real potential for influencing ALJs to rule in accordance with what is perceived to be likely to bring about good evaluations and thereby to protect their present and future job security, even when rulings based solely on the facts and law of cases would reach different results. While this may not happen in such a blatant manner as the example given in a Dilbert cartoon, in which the boss evaluated his employees on the basis of how many boxes of Girl Scout cookies they bought from his daughter, it can happen, on many levels.

When such compromised independence in decision-making occurs, it does not necessarily occur through direct attempts to influence outcomes. It will, in fact, generally occur in more subtle ways: For example, an ALJ knows the chief judge -- or some other evaluator -- strongly holds a particular position on a legal issue and feels that anyone who disagrees is wrong; the ALJ disagrees, has a good legal rationale to back up his or her viewpoint, and rules in the way he or she interprets the law when the issue comes up in a proceeding; the evaluator may be tempted to mark the ALJ down in a category such as "Applies law to facts appropriately," and may do so.

Another possible example of undue influence that can manifest itself at evaluation time arises from the fact that most judges have cases, more or less frequently, in which a ruling in one or another way will be controversial, possibly have great financial and human

consequences, and be apt to provoke strong negative reactions from some individual or segment of the community, or from government agency personnel. Ruling the way the law, and the conscience of the judge, dictate, may prove costly, not only for the state or other party, but also for the judge, and other judges and co-workers -- for example, when budgets are under consideration and those offended by a ruling may include some who may have influence in the budget approval or cutting process.

A chief judge may have difficulty addressing pressure that may be exerted, and may in turn exert pressure on judges who make such unpopular rulings. This may manifest itself in the evaluation process, no matter how much care is taken to prevent this from happening by, for example, limiting categories for evaluation to ones which do not address outcomes in cases. A category such as "maintains harmonious relations with the public, agencies and co-workers," could, for example, be subject to the subtle exertion of improper influence, based on the negative reactions the judge's decision might produce, both in members of the public and other agencies, and in the chief judge. Such things occur, human nature being what it is, and judges not being exempt from the basic principles of human nature.<sup>92</sup>

As Kenneth K. Stuart, Chief Judge of the 18th Judicial District of Colorado and a participant in a 1995 American Judicature Society panel presentation on "Evaluating the performance of judges standing for retention," has noted, such human reactions can occur in the judicial branch as well. Discussing the strengths and weaknesses in Colorado's judicial evaluation program, Judge Stuart noted several strengths in the system, including judicial improvement and bringing into the system a

---

<sup>92</sup>Other examples of how independent decision-making can be compromised in subtle ways were discussed at the 1996 Annual Meeting and Conference of the NAALJ, held in Nashville, Tennessee, in November, 1996, in a panel discussion on the subject of "Due Process, Ex Parte Communications, and the Tension Between Independent Decision-Making and Administrative Concerns." (Copies of the hypothetical examples and other materials used in this discussion on file with the NAALJ Journal at Loyola University Chicago School of Law, and are also available from the author.)

It should be noted that such problematic situations may be improved, but are not necessarily eliminated, in central panels of ALJs. While there is increased separation of ALJs from agency personnel, agencies can still have an impact on budgets, for example. ALJs may therefore be subjected to pressures from not just one agency, but many, in the absence of appropriate structural, procedural, and budgetary protections against such pressures.

large cross-section of people, and also noted some weaknesses:

Another area that some people may view as a strength -- others view it as a weakness -- is that because you are getting the feedback it does tend to temper judicial arrogance. A weakness though, is that it can -- *it's not just a philosophical concern, but a real concern* -- affect judicial independence because especially near retention time judges become very worried about the way the [evaluation] commission is going to react.<sup>93</sup>

---

<sup>93</sup>Transcript of panel presentation, *Evaluating the Performance of Judges Standing for Retention*, 79 JUDICATURE 190, 194 (Jan.-Feb. 1996) (emphasis added).

At the 1996 meeting of the American Judicature Society, a panel discussion was held on the subject of "What is Judicial Independence?" During this discussion, in which it was noted that "never in recent memory has there been more discussion about judicial independence than in the last year," an audience member offered a comment, which initiated a short exchange of views that illustrates well some of the widely ranging views on the subject of judicial independence:

I'm a federal judge in Detroit. I assure you that criticism of an action on a suppression motion by the chairperson of the Senate Judiciary Committee has a powerful effect on many federal district judges who aspire to the court of appeals. Federal judges at the district court level, particularly younger ones, aspire to the court of appeals, and they know their votes are being watched, their decisions are being analyzed. Young judges on the courts of appeals aspire to the Supreme Court of the United States. And I assure you, they think their votes are being watched. Now whether that's good or bad in terms of judicial independence, I can't tell you. But it's simply a fact of life. Everybody wants to be loved.

*Id.* Panel moderator Erwin Chemerinsky, Legion Lex Professor at University of Southern California Law Center, responded, "Why believe that is bad rather than good? When we know we're being watched, maybe we're more careful." This prompted panel member Kenneth C. Jenne II, Florida State Democratic leader, to state:

To me, the judge's comment is the most frightening commentary on appointed judges if this ambition is so overwhelming that they are going to orient their decisions and actions to please one or two people. I thought a judge was an individual who was not faint-hearted. My intellectual description of a judge is someone with backbone and courage who's going to do the right thing. When he or she takes an oath of office, it is meaningful. There has to be a certain amount of courage in public office, and if one's ambition is thwarted because of a ruling, I think that individual would make a poor judge to begin with. My concept of whom I want to appear before is not a faint-hearted judge who's going to be afraid to rule. I would hope every federal judge would answer like Judge [William M.] Hoeveler [U. S. District Court Judge for the Southern District of Florida], that they pay attention to criticism and then they do what they want to do.

Edited transcript, 80 JUDICATURE 73, 81 (Sept.-Oct. 1996).

Judge Stuart gave another example of how unsubstantiated complaints that make their way into the evaluation process can affect judges, describing a criticism leveled at a judge that he had exhibited gender bias, with no statistical evidence to prove it and no way for the judge to respond, because the input received by the evaluation commission was confidential. Despite the lack of evidence, and what “the judges feel to a certain extent [to be] a lack of process,” the judge in question lost his retention election, and “[t]he general sense [was] that that one criticism cost him the election.”<sup>94</sup>

Judge Stuart also described the way in which other forms of public perceptions and subtle pressures can influence evaluations of judges in the judicial branch, describing what he termed the “first weakness” of the system:

There’s a public perception of government inefficiency and bureaucracy that spreads over to the judiciary, and there is some pressure on commissions to buy into that feeding frenzy and recommend against a particular judge. If you are always saying good things about judges and always recommending them for retention, in the eyes of the public your credibility goes down. So, there’s a sense among the commissions informally (*no one would say this publicly*) that there have not been enough negative recommendations. This, obviously, is of great concern to judges.<sup>95</sup>

Going on to describe information he had received from a statistician who was analyzing questionnaire results, to the effect that juror questionnaires were unanimously favorable to judges, Judge Stuart noted that the juror results were discounted because they were all

---

<sup>94</sup>See transcript of 1995 panel presentation, *supra* note 93, at 195.

<sup>95</sup>*Id.* at 193 (emphasis added). See also Jean Tirole, *Collusion and the Theory of Organizations*, in *ADVANCES IN ECONOMIC THEORY: SIXTH WORLD CONGRESS* vol. 2 at 151, (Jean-Jacques Laffont ed. 1992). Analyzing the phenomenon of collusion in groups, Tirole notes that “behavior is often best predicted by the analysis of group as well as individual incentives,” and shows that “incentive structures must account for the possibility that members collude to manipulate their functioning.”

positive.<sup>96</sup>

If such questions can be raised when evaluations of judicial branch judges are performed by commissions, which are generally made up of groups of persons appointed because they are highly respected in the legal community, is not the possibility of such subtle compromising of the process even more likely to arise when there is but a single evaluator, or one primary evaluator? Certainly, while a “herd” instinct may arise in groups, it is likely that individual action, not so exposed to the view of others, may be even more easily corrupted by inappropriate influences.

It should be recognized that, while inappropriate influence and coercion in evaluations may be conscious and blatant, they are more likely to be subtle, and possibly semi-conscious or unconscious. Both evaluators and those evaluated partake of such common human imperfections as limited awareness of differences, unacknowledged and/or unconscious biases and prejudices,<sup>97</sup> and inconsistent moods that may affect perceptions, actions, and decision-making. Separately or combined, these kinds of factors can lead to inappropriate influence, and compromise decision-making independence.

When this occurs, and when there are pronounced discrepancies between what is likely to bring about good evaluations, and what actually constitutes appropriate performance and decision-making based on the facts and law of cases, the integrity of the administrative adjudicative system is seriously compromised. Even minimal and subtle compromise of this sort renders due process more a sham than a reality; and a sham may be said to be worse than an outright, blatant compromise of principle: because it *looks* real, it deceives and lulls into a sense of false security.<sup>98</sup>

This danger must not be ignored, if real fairness and neutrality are to be desired and achieved on a consistent basis, rather than randomly,

---

<sup>96</sup>Transcript of 1995 panel presentation, *supra* note 93, at 193. At another point in the discussion, Judge Stuart and Michael D. Zimmerman, chief justice of the Utah Supreme Court, observed that judicial temperament and demeanor, and “arbitrary setting of dates, decision-making -- that sort of thing” led to more negative evaluations than other characteristics such as legal knowledge and ability. *Id.* at 197.

<sup>97</sup> See *supra*, text accompanying notes 40-47 and 61-86.

<sup>98</sup>As Redish and Marshall observe, “It is dangerous to create an illusion of predictability when the decision was in fact reached on the basis of irrational factors.” Redish, *supra* note 47 at 486.



by the mere happenstance of whether in any given case there is some pressure, invisible to one or more parties in the case, influencing the ALJ to make rulings differently than she or he would rule based solely on the facts and law of the case.<sup>99</sup> All of the examples given above could obviously intimidate judges into ruling in ways that are likely to elicit more positive responses from evaluators, especially in controversial cases, and especially in cases in which the facts and the law are fairly evenly balanced and the ruling could go either way -- why not choose the easy way, when the stakes are so high from a personal standpoint (in which one's career could be, if not "decimated," then at least negatively affected)?

In such situations judges must take special care not to lean either way in response to such perceived pressure, and this may take courage, and some amount of struggle to find the right way to proceed under the responsibility of true independent decision-making. And yet, as Professor Burbank reminds us, we should not "increase the number of occasions when personal courage is necessary,"<sup>100</sup> through, for example, management and evaluation practices in an office.

Doing so may have broader impact than might be expected, because the potential for inappropriate influence extends beyond the evaluated judge who has a difficult decision to make. Other judges who observe negative consequences for a judge in a difficult position who acts with courage may be even more intimidated when they find themselves in a difficult position, and if the first judge takes the easy way and thereby avoids "punishment," or is even rewarded, others may become more likely to follow suit. An office or judicial system can be infected with the subtle virus of trying to please, and the threat of public disapprobation will make it "extremely difficult" to go against the prevailing winds.<sup>101</sup>

As Princeton Professor B. Douglas Bernheim has recognized, "[w]hen status is sufficiently important relative to intrinsic utility, many

---

<sup>99</sup>As Judge Larkin observes, "No judge's individual decision should be foreordained because of hidden pressures or unarticulated premises." Larkin, *supra*, note 26, at 5.

<sup>100</sup>Burbank, *supra* note 20, at 120.

<sup>101</sup>In Judge Larkin's words, "[t]he more ominous threats to judicial independence are more subtle, stealthy, and subliminal. They involve forces that operate on individual judges as well as entire court systems. Sometimes these influences are overt. More often they are silent and secretive, but no less telling." Larkin, *supra*, note 26, at 7. See also note 18, *supra*.

individuals conform to a single, homogeneous standard of behavior, despite heterogeneous underlying preferences. They are willing to conform because they recognize that even small departures from the social norm will seriously impair their status.”<sup>102</sup>

Those who believe judges are immune from such pressures to conform may be wearing blinders unaware. That such a phenomenon can occur without being acknowledged is actually more likely than that it would be acknowledged, in a sense. As illustrated by any number of recent situations in which public figures have gotten into trouble for doing what had always previously been accepted on such a “see-no-evil” basis (political campaign financing is but one example) inappropriate activity can occur on a broad, generally accepted, but unspoken-of, basis. While judges may be above some of this by virtue of the countervailing pressure to comply with the Code of Judicial Conduct, it is extremely unlikely that as a group they are completely able to withstand pressure to conform, or “go along to get along,” especially when it is the easier thing to do, and when not to go along can have extreme negative personal and financial consequences, and negatively affect one’s status.

It has been said that the “greatest threat to the rule of law comes from those judges who . . . make compromises in order to stay in office . . . .”<sup>103</sup> Good selection criteria and training programs in ethics for judges can help to address this problem. However, it must be recognized that judges are human beings, and that only by addressing such problems from a systemic perspective will there be broadly effective remedies for such compromise of principles. Evaluation programs and their effects on independence must be taken into account in such approaches.

## 2. Effects on Quality, Efficiency, Morale, and Productivity

In addition to the danger of compromising independent decision-making, other, related negative consequences of traditional evaluation systems include: competition rather than cooperation among ALJs (for

---

<sup>102</sup>B. Douglas Bernheim, *A Theory of Conformity*, 102 J. POL. ECON. 841 (1994).

<sup>103</sup>Stephen B. Bright, *Political Attacks on the Judiciary*, 80 JUDICATURE 165, 172 (Jan.-Feb. 1997).

example, in sharing resources and helpful feedback), which in turn causes quality of work product to suffer; the discouragement of independent and creative thinking, which causes overall quality and efficiency of the workplace to suffer; and lowered morale, which causes decreases in productivity and efficiency.

Given that ethical ex parte prohibitions appropriately, but severely, limit a judge's freedom to talk about cases with pretty much everyone except other judges,<sup>104</sup> the importance of maintaining a healthy

---

<sup>104</sup>Canon 3B(7) of the 1990 ABA MODEL CODE OF JUDICIAL CONDUCT provides as follows:

(7) A judge shall accord to every person who has a legal interest in a proceeding, or that person's lawyer, the right to be heard according to law. A judge shall not initiate, permit, or consider ex parte communications, or consider other communications made to the judge outside the presence of the parties concerning a pending or impending proceeding except that:

(a) Where circumstances require, ex parte communications for scheduling, administrative purposes or emergencies that do not deal with substantive matters or issues on the merits are authorized, provided:

(i) the judge reasonably believes that no party will gain a procedural or tactical advantage as a result of the ex parte communication, and

(ii) the judge makes provision promptly to notify all other parties of the substance of the ex parte communication and allows an opportunity to respond.

(b) A judge may obtain the advice of a disinterested expert on the law applicable to a proceeding before the judge if the judge gives notice to the parties of the person consulted and the substance of the advice, and affords the parties reasonable opportunity to respond.

(c) A judge may consult with court personnel whose function is to aid the judge in carrying out the judge's adjudicative responsibilities or with other judges.

(d) A judge may, with the consent of the parties, confer separately with the parties and their lawyers in an effort to mediate or settle matters pending before the judge.

(e) A judge may initiate or consider any ex parte communications when expressly authorized by law to do so.

*Commentary:*

*The proscription against communications concerning a proceeding includes communications from lawyers, law teachers, and other persons who are not participants in the proceeding, except to the limited extent permitted.*

*To the extent reasonably possible, all parties or their lawyers shall be included in communications with a judge.*

*Whenever presence of a party or notice to a party is required by Section 3B(7), it is the party's lawyer, or if the party is unrepresented the party, who is to be present or to whom notice is to be given.*

*An appropriate and often desirable procedure for a court to obtain the advice of a disinterested expert on legal issues is to invite the expert to file a brief amicus curiae.*

*Certain ex parte communication is approved by Section 3B(7) to facilitate scheduling and other administrative purposes and to*

collegiality among judges in the same office cannot be gainsaid. When this is lost, not only does lowered morale decrease efficiency and productivity both individually and on an officewide level, accuracy may also be lost, especially where many cases involving the same issues are heard, and shared information about the latest appellate rulings and statutory and regulatory amendments can be important, for example, in order to avoid costly remands.

In addition, although the discouragement of independent and creative thinking may have short-term benefits such as more compliant employees and reduced conflict on an open level, the long-term results of such an approach will be a low-functioning workplace, of the sort that leads to various unfortunate results, ranging from well-known jokes about dull-witted government employees, to compromised service to the public, including those citizens who are ill-equipped on their own to address legal improprieties on the part of judges. All of us, citizens and government workers alike, deserve better than this.

### C. Human Sources of Negative Consequences and the Need for Alternative Approaches

Common sense and anecdotal evidence illustrate how traditional evaluation systems may have unintended negative consequences. For

---

*accommodate emergencies. In general, however, a judge must discourage ex parte communication and allow it only if all the criteria stated in Section 3B(7) are clearly met. A judge must disclose to all parties all ex parte communications described in Sections 3B(7)(a) and 3B(7)(b) regarding a proceeding pending or impending before the judge.*

*A judge must not independently investigate facts in a case and must consider only the evidence presented.*

*A judge may request a party to submit proposed findings of fact and conclusions of law, so long as the other parties are apprised of the request and are given an opportunity to respond to the proposed findings and conclusions.*

*A judge must make reasonable efforts, including the provision of appropriate supervision, to ensure that Section 3B(7) is not violated through law clerks or other personnel on the judge's staff.*

*If communication between the trial judge and the appellate court with respect to a proceeding is permitted, a copy of any written communication or the substance of any oral communication should be provided to all parties.*

example, more than a few ALJs (along with employees of every type) have observed the sense of unease that creeps into and pervades an office at evaluation time each year, quarter, etc. It might be asked why adult professionals who have no ostensible reason to fear evaluation react in such a way.

Perhaps it is because they, as human beings, recognize that, however lofty the goals and rhetoric of any system, those who administer it are, as human beings, subject to such human failings as inconsistent and whimsical moods, limited awareness of differences, and even biases and prejudices, be they conscious or unconscious. It is implicitly understood that any evaluation or ranking of ALJ performance is essentially unilateral, not subject to the procedural protections of the legal system, and therefore more subject to human bias and error. Also, in this time of well-publicized government cutbacks and "job anxiety," people are generally aware that traditional evaluations may later be used in ways that may adversely affect their livelihoods and careers, even without fault in the usual sense, and with relatively little meaningful opportunity to respond effectively to the contents of such evaluations.

Harking back again to Hamilton, and recalling what Prendergast et al. have shown us, we can see that the negative unintended consequences of traditional performance evaluation systems become more serious, the less job security an ALJ enjoys, and the more his or her financial well-being may be affected by evaluations. While many ALJs enjoy full civil service protections and are removable only for cause, others are removable at will, have no job security other than the good will of their employer, and have only those protections under the law that apply to any at-will employee. The less assurance of security regarding discipline and removal procedures and standards, and the more the livelihoods of ALJs are affected by evaluations, the more important it is to address and avoid the potential negative consequences of an evaluation system that may compromise independent decision-making, through inadvertently pressuring ALJs to decide cases based, however subtly, on inappropriate factors.

These are sensitive issues, because everyone likes to think that she or he is fair, probably especially those of us involved in various forms of adjudicating disputes for a living. However, it must be remembered

that the real issue, in any serious discussion of evaluations and independent decision-making, is not whether individual ALJs or evaluators are capable of being fair, but how weaknesses in a system can threaten to compromise even the best of us in potential worst-case scenarios.

Traditional evaluation systems purport to improve performance, a legitimate concern: few if any would disagree that judges should strive to improve their level of competence in all areas of judicial performance, or that judges who are truly incompetent and incapable of improving their level of competence should be removed, or at least encouraged to leave. However, fears that, without a traditional performance evaluation system, traditionally implemented, some weak ALJs will “fall through the cracks” and compromise quality, may be overstated and misplaced. In the judicial branch, strict limitations on the oversight and removal of judges have traditionally been deemed to be a risk worth taking, based on the recognition that too much oversight and/or too easy removal will lead to all-too-human compromising of judicial independence, and based on the conviction that the strong values of due process<sup>105</sup> outweigh the danger that some judges will be less than maximally competent.<sup>106</sup>

For the same reasons, concerns about competence in the administrative judiciary may be better addressed through alternative approaches, which are grounded not only in due process values but also in sound management principles. Alternative approaches to structuring and implementing standards-setting and enforcement systems for ALJs may also better address and prevent the potential negative consequences of traditional evaluation models.

---

<sup>105</sup>To paraphrase Judge Irving R. Kaufman, judicial independence is “not . . . for the benefit of the judges, but for the benefit of the judged.” Kaufman, *supra* note 37, at 690, citing Kurland, *The Constitution and the Tenure of Federal Judges: Some Notes from History*, 36 U. CHI. L. REV. 665, 698 (1969).

<sup>106</sup>As Kaufman observed (after suggesting that peer pressure is a “potent tool” for addressing judicial incompetence), “Our judicial system can better survive the much discussed but rarely existent senile or inebriate judge than it can withstand the loss of judicial independence that would ensue if removal of judges could be effected by a procedure too facile or a standard too malleable.” *Id.* at 709, 715-716. Although Kaufman was speaking of the federal judiciary, it is suggested that his observations apply equally to administrative law judges, as long as the role of ALJ remains one of implementing, in a judicial model, due process in proceedings affecting the interests of persons before them.

#### IV. Proposal: Implementing Fundamental Principles in Traditional Settings

The understanding that can be gained from the field of management, on how independent thinking, efficiency, and productivity can be compromised by traditional performance evaluation programs in any setting, has increased relevance and urgency in legal systems, in which the fundamental principles of due process and independent decision-making, and the perception as well as the reality of fairness, are especially to be valued. Alternative approaches that encourage independent thinking and truthful, constructive communication may more effectively improve performance, while at the same time fostering, rather than hampering, independent decision-making, thereby fulfilling the values of due process.

Whether such alternative approaches are arrived at through creation "from the ground up," or, as will be the case with most ALJ offices, through the restructuring or fine-tuning of an existing system, a careful analysis must be made of all potential weak points, and such weak points must be reinforced within the context of an effective overall structure, if an optimally performing system is desired. All too often, little attention is paid to the actual structure of a system, and heavy reliance is placed on the abilities, fairness, and good will of the individuals who administer it. A more carefully structured system is more likely, however, to elicit the best from the people in it, as Deming recognized. Conversely, a hastily or haphazardly structured system is more likely to fail as a result of human error, failure of good will, or lack of fortitude and clear vision.

The administrative law system is not exempt from these somewhat self-evident principles. While ALJ decisions may be subject to more levels of review than are judicial branch trial judge decisions, and while ALJ rulings may take into account agency policy -- as found in administrative rules, administrative adjudicative precedent, and official and executive orders -- to a degree not generally found in the court system, it should be remembered that the central function of ALJs is to fulfill the requirements of due process. Their rulings should therefore be based on the same time-tested principles of the American legal

system, and made with the same responsibility to independent decision-making, as that which should be borne by judges and others performing judicial functions in the judicial system. And any management system for ALJs should foster such responsibility to independent decision-making, in order to fulfill more effectively the values of due process.

To achieve this end, an appropriate standards-setting and enforcement program for ALJs should incorporate rigorous selection criteria and procedures, high standards of conduct, effective continuing education and training, appropriate and meaningful observation and feedback for self-improvement, an effective and fair complaint system, and clearly-defined and appropriate discipline and removal procedures.

I propose a system that:

- \* provides for careful evaluation of ALJ applicants on the front end, based on qualifications, experience, and character;
- \* incorporates standards of performance from codes of judicial conduct, with any other standards developed cautiously, with careful attention to the need to protect independent decision-making, and with full participation by all judges who will be affected;
- \* includes workable and useful methods of training, observation, and feedback for self-improvement, which avoid any compromise of independent decision-making as much as possible;
- \* assures both ease of making complaints, and fair evaluation of complaints against ALJs who have allegedly violated appropriate standards of conduct, or who are allegedly incompetent; and
- \* protects the rights of ALJs to be disciplined and removed only in accordance with clearly defined, appropriate, and fair standards and procedures, based on the same principles of due process and independent decision-making that should be applied by ALJs in the cases they hear. Any reduction-in-force procedures should also be clearly defined and based on neutral, value-free criteria such as seniority, with appropriate provision for any groups who may have been historically underrepresented in an agency or ALJ office, in accordance with relevant law.<sup>107</sup>

---

<sup>107</sup>The Model Act Creating a State Central Hearing Agency (Office of Administrative Hearings), adopted by the American Bar Association House of Delegates at the mid-year meeting in San Antonio on February 3, 1997, provides, at section 1-6(a)(4), that ALJs are to



It must be recognized that the context for implementation of these goals, whether in a newly-created system or in modifying an existing system, will always be the real world we live in. Human organizations do not operate in a vacuum, and the real-world setting for any adjudicatory system must be acknowledged and analyzed, both for its potential weak points, and for its strong points. Trouble-shooting should be an ongoing and welcome activity, if an organization is to be effective. In addition, strengths should be recognized and utilized, and any temptation to shun the old and the traditional, and to glorify the innovative and the new merely because it is new, should be strenuously fought and overcome. Similarly, any tendency to strike a radical or extreme pose for political reasons (“ask for more than you want to get what you want”) should be resisted, at least in the context of the actual structuring of a system.

The structuring or restructuring of any standards-setting and enforcement system for ALJs should, rather, be undertaken in a spirit of moderation, constructiveness, and inclusiveness, while at the same time subjecting all parts thereof to the strict scrutiny of whether they foster or hamper due process, and whether they are in accordance with

---

be “removed, suspended, demoted, or subject to disciplinary or adverse actions, including any action that might later influence a reduction in force, only for good cause, after notice and an opportunity to be heard in an Administrative Procedure Act or other statutory-type hearing and a finding of good cause by an impartial hearing officer.” The Act also provides, at section 1-6(a)(5), that ALJs are to “be subject to a reduction in force only in accordance with established, objective civil service or merit system procedures.” The words in section 1-6(a)(4), “including any action that might later influence a reduction in force,” arrived at through a compromise between ALJs present at a February 1, 1997, meeting of the Administrative and Regulatory Law Section of the ABA and Columbia Law Professor Peter Strauss, were intended to allow an ALJ to receive a hearing on any action, including any evaluation or any other determination that might influence a reduction in force, when such action is taken, rather than after the fact, when the reduction in force actually occurs. The concern of Professor Strauss was that managers not be held up with hearings at the time of any reduction in force. However, all agreed that, whenever any action occurs, whether it be an evaluation or some other determination related to performance, of a sort that might somehow be used by the manager to decide who will be laid off in a reduction in force, the ALJ subject to such action would have the right to a full hearing under section 1-6(a)(4) at the time of such evaluation or other determination or action, to contest the determination(s) made in such action.

As of April, 1997, at least one state, Alaska, was considering adopting the Model Act. House Bill 232, Twentieth Legislature - First Session, State of Alaska, introduced 4/4/97. (Copy on file in the offices of the NAALJ Journal, at Loyola University of Chicago School of Law.)

sound management principles that further independent and creative thinking, quality, efficiency, morale, and productivity.

### A. Standards for Selection of Administrative Law Judges

ALJ selection criteria should be rigorous, and should include careful evaluation of qualifications, experience, and character, to assure the selection of highly-qualified, competent ALJs who have a broad awareness of differences and are able to apply the law in a fair and neutral manner. The procedures for such selection should be crafted to assure fairness, and the selection of the best candidates. As noted by Professor Burbank, inappropriate criteria for the selection of judges can “abort” the issue of judicial independence on the front end.<sup>108</sup> Independent screening panels such as are used in the selection of judges in the judicial branch, and of ALJs in some states, may be used, both to ensure quality more effectively, and to avoid inappropriate political or other influence in the selection of ALJs, as much as possible.

### B. Standards of Conduct for Administrative Law Judges

The most important standards of conduct for ALJs, as for all judges, are those set forth in the various judicial codes of conduct, and these should be incorporated and applied to ALJs.<sup>109</sup> Such codes should

---

<sup>108</sup>Burbank states as follows:

[T]aken to extremes, the use of criteria for appointment to the federal bench that are tied to an individual’s views on important issues of law and social policy can fairly be seen as an attempt to abort, rather than to predict, the issue of judicial independence. That is because true judicial independence not only requires insulation from external forces that would seek to influence *judicial* judgment other than through the judicial process. It also requires insulation from those forces, external and internal, that so constrain *human* judgment as to subvert the judicial process.

Burbank, *supra* note 20, at 120. (Emphasis in original.) The same principle would seem to apply to the selection of ALJs.

<sup>109</sup>It is noted that this is already done in many ALJ offices, especially central panels of state ALJs, and that there are various model codes of judicial conduct for ALJs in existence, including codes adopted by the NAALJ and the ABA National Conference of Administrative Law Judges, and the Model Code of Conduct for Administrative Law Judges of State Central Panels. The first two of these codes are presently under revision to incorporate appropriate changes based on the 1990 ABA Model Code of Judicial Conduct. See, e.g., *supra* note 104.

make use of the most recent, well-thought-out model for judicial conduct, the 1990 ABA Model Code of Judicial Conduct. This code may be adapted to the particular and unique circumstances of ALJs as necessary, but this should be done sparingly, to assure high standards of conduct and professionalism, and to maintain judicial independence. All ALJs should be thoroughly trained in the ethics of being a judge, in the subtleties and fine points of all requirements of relevant codes of conduct, and in the need to conscientiously adhere to all such requirements. Compliance with the judicial code of conduct should be emphasized, and continuing education and training should be provided in the area of ethics and appropriate judicial conduct for all judges, on a regular basis.

This article is not intended to encompass an exhaustive study or analysis of all possible standards of conduct that might be included in any code of conduct, or in any necessary or possible evaluation instruments regarding ALJs. There are a multitude of sources on particular criteria for evaluation, including the American Bar Association and the National Center for State Courts. In order to protect fair and independent decision-making, reference should be made to these in designing any evaluation standards for ALJs, rather than to traditional models for non-judicial employees. The use of any existing models of evaluations for ALJs should be limited to standards that incorporate a judicial model, and that meet the test of reasonableness from a judicial perspective. The reason for this is not to enhance the prestige of the ALJ position, but to enhance the professionalism and independence of ALJs, not for their benefit so much as for the benefit of the parties who come before them (both the state and private parties).

With this said, I would offer a few considerations to take into account in drawing up standards of conduct for ALJs. First, while it has been suggested that overly broad and "amorphous" standards of conduct may impinge on judicial independence,<sup>110</sup> overly detailed and extensive standards of conduct for evaluation purposes may have the same result, by in effect micro-managing judicial performance, a

---

The third has been formally adopted in at least one state with a central panel of ALJs, Colorado. COLO. REV. STAT. §24-30-1003(4), 1 COLO. CODE REGS. §104-2 (1994).

<sup>110</sup>James P. Timony, *Disciplinary Proceedings Against Federal Administrative Law Judges*, 6 W. NEW ENG. L. REV. 807, 828 (1984).

function which by its very nature demands a high degree of autonomy. While reasonable standards of performance and productivity should be expected of ALJs, the key word in this regard is "reasonable." Such means as minutely detailed descriptions of expected performance can lead to lowered morale and loss of pride in one's work, and undermine effective and independent decision-making. As indicated in the section below on ongoing observation and feedback on judicial performance, there are alternative methods of improving performance that are more useful.

An example of standards that might be adapted for use with ALJs is found in the ABA Guidelines for the Evaluation of Judicial Performance. The ABA criteria include:

- (1) Integrity -- avoidance of impropriety and appearance of impropriety, freedom from bias, impartiality;
- (2) Knowledge and understanding of the law -- legally sound decisions, knowledge of substantive, procedural and evidentiary law of the jurisdiction, proper application of judicial precedent;
- (3) Communication skills -- clarity of bench rulings and other oral communications, quality of written opinions, sensitivity to the impact of demeanor and other non-verbal communications;
- (4) Preparation, attentiveness and control over proceedings -- courtesy to all parties, willingness to allow legally interested persons to be heard unless precluded by law;
- (5) Managerial skills -- devoting appropriate time to pending matters, discharging administrative responsibilities diligently;
- (6) Punctuality -- prompt disposition of pending matters and meeting commitments of time according to rules of court;
- (7) Service to the profession -- attendance at and participation in continuing legal education, ensuring that the court is serving the public to the best of its ability;
- (8) Effectiveness in working with other judges -- extending ideas and opinions when on multi-judge panel, soundly critiquing work of colleagues.<sup>111</sup>

---

<sup>111</sup>ABA Special Comm. on Evaluation of Judicial Performance, Guidelines for the Evaluation of Judicial Performance, (1985), as quoted in Lubbers, *supra* note 6, at 608.

Perhaps the most important principle to adhere to in defining standards of judicial performance for evaluation purposes is that most any standards, including those listed above, can compromise independence if approached from a biased (consciously or unconsciously), punitive, or threatening point of view, as illustrated in the section of this article on unintended consequences of traditional evaluation systems. The best approach to take in creating any necessary evaluation standards of conduct, over and above those set forth in relevant codes of judicial conduct, is therefore a cautious one. Each standard that is considered for inclusion should be questioned: Is evaluation of a judge by this standard truly necessary in order to achieve good judicial performance? Can compliance with this standard be accomplished more effectively by means other than evaluation? Could this standard compromise or subvert, even minimally or subtly, the responsibility of ALJs to decide cases independently, based solely on the facts and law of the cases? How might this standard be used by a weak evaluator, in a worst-case scenario, in a way that might compromise judicial independence or performance? While perfect answers may not be attained, it is important to ask such questions in good faith, if an optimal program is to be designed.

The importance of the participation of those to be evaluated in designing evaluation standards is also important, and participants “should be given the opportunity to vent frustrations, comment, offer suggestions, and review the work product as it takes shape. . . . If this open approach is not adopted and maintained, the judicial performance evaluation program may make judges [and other affected persons] fearful, and the informal network will become active, [resulting] in rumors and half-truths [and] institution of a less effective program . . .”<sup>112</sup>

The development of standards of conduct should be undertaken by ALJs themselves, or by committees or commissions made up primarily of attorneys and ALJs who understand the potential negative consequences of evaluation programs to judicial independence. Inappropriate input or influence from parties, agency personnel, or others with case outcome-related interests should be avoided, and any

---

<sup>112</sup>See Farthing-Capowich, *supra* note 9, at 24.

input from such persons would best be limited to contributing suggestions for portions of any survey instruments that may be developed for use with groups of such persons. Simplicity, straightfowardness, and clarity should be striven for in defining any standards and questions for any surveys that may be used, in order to avoid ambiguity and confusion, and to give those who will be evaluated adequate notice of expected standards of conduct.

### C. Continuing Education, Training, and Professional Development

Continuing education, training, and professional development activities for ALJs should be provided, and attendance should be encouraged, and mandated at an appropriate minimal level, in ethics and all areas of skills improvement: procedural rules, judicial demeanor and basics of conducting hearings, evidence, case management procedures and skills, legal writing, and alternative dispute resolution, to name just a few. Where weakness is perceived in an ALJ's performance in a particular area, extra educational, training, and/or professional development activities in that area should be provided for that judge. In order to minimize costs, ALJs should be encouraged to share their own knowledge in their areas of expertise with each other, both informally and in formal training sessions. ALJs should be encouraged to join professional organizations through which additional education, training, professional development, and collegial sharing can be obtained.

It must be recognized that, if accountability, quality, productivity, efficiency, and fairness are to be expected of ALJs, then an investment must be made in the administrative judiciary. Any good professional continues to learn and keep abreast of current developments in a field of endeavor, and good ALJs should maintain their familiarity with current developments in the law relevant to their particular areas of practice, along with the more general areas of ethics, professional responsibility, and skills development. In this regard, it is important to keep in mind a lesson learned from the field of education: Learners who are respected and expected to do well are more likely to do well than are students for whom teachers have lower expectations. Continuing educational, training, and professional development

activities should be encouraged as an enjoyable and stimulating aspect of professionalism, rather than as an arduous requirement.

In developing a judicial education program, the principles of experiential learning theory and learning styles developed by Professor David A. Kolb, of Case Western Reserve University, and Patricia H. Murrell and Charles S. Claxton, of the Leadership Institute in Judicial Education, should be studied and applied, in order to maximize learning and its real-world application. These include the idea that learning occurs as a melding of two dimensions, “prehending” or taking in information, and transforming or processing information, and the fact that people have different styles of learning in both these dimensions: Some people naturally prefer to take in learning based on concrete experience, whereas others prefer abstract conceptualization. Some people process information better by doing, others by watching and reflecting. These preferences, in various combinations, produce four different learning styles: “divergers,” who prefer concrete and reflective learning; “assimilators,” who prefer abstract and reflective learning; “convergers,” who prefer abstract and active learning; and “accommodators,” who prefer concrete and active learning. Including activities in all modes -- direct concrete experience, reflection on experience, abstract conceptualization, and application of abstract principles -- will make any learning experience more successful for more people, including judges.<sup>113</sup>

Professional development activities may also include private reading and research projects, and self-evaluation exercises based on

---

<sup>113</sup>Information learned at condensed session on Leadership Through Judicial Education, presented by Patricia H. Murrell, September 27, 1996, at the National Association of Women Judges Conference in Memphis, Tennessee. See Charles S. Claxton and Patricia H. Murrell, *Education for Development: Principles and Practices in Judicial Education*, Leadership Institute in Judicial Education, JERITT Monograph Three, 1992. (Copy on file in the offices of the NAALJ Journal, at Loyola University of Chicago School of Law.) In their monograph, Claxton and Murrell also discuss different stages of thinking -- dualistic, contextual, and integrated -- and the qualities of “highly developed judges.” *Id.* at 38 et seq. More information on these concepts can be obtained from Patricia H. Murrell, Center for Study of Higher Education, Education Annex One, The University of Memphis, Memphis, Tennessee 38152. See also Gina L. Hale, *Professionalism: A Call to Excellence*, XVI J. NAALJ 5, 13 (Spring 1996), the 1995 NAALJ Fellowship paper, in which the author, citing among others Claxton and Murrell, advocates the idea of developing “learning organizations,” which “perceive and foster learning as an integral part of the organizational culture.”

models of desired conduct or achievement that are arrived at cooperatively with those who will be affected.

#### D. Ongoing Observation and Feedback on Judicial Performance

Any ongoing, regular formalized methods of observing and providing feedback on the performance of ALJs should be carefully considered both from a legal and a management standpoint, as elucidated by the innovative management research and theory discussed in this article, in order to avoid compromising the duty and responsibility of ALJs to decide cases independently.

It should be recognized, as Professor Prendergast has pointed out, that performance evaluation is a process by which “humans judge other humans,”<sup>114</sup> and place “value” on them as functioning members of society. And yet, it should be acknowledged, traditional performance evaluations violate some of the basic principles of judging in the American legal system, in that they are essentially unilateral, they are based largely on subjective impressions rather than evidence, and they provide little opportunity to respond effectively. Unfortunately, efforts to avoid subjectivity can backfire. For example, when case load and other statistics are used in evaluations, this can lead to distortion and compromised impartiality in decision-making,<sup>115</sup> and it should not be forgotten that statistics can be manipulated and interpreted in any number of ways, as any reasonably aware reader of the popular press knows.<sup>116</sup> Finally, although most traditional evaluation forms include a space for the evaluated employee to write a response, this will generally not lead to any revision of the statements made in the

---

<sup>114</sup>Prendergast, *supra* note 62 at 355.

<sup>115</sup>See O’Keeffe, *supra*, note 9, at 621: “. . . [E]ven if production standards do not affect the substance of a case directly, they may do so indirectly through the changes they engender in the decision-making process.” See also *supra*, text accompanying notes 84-85.

<sup>116</sup>Comparing numbers of cases or orders can obviously be deceptive, since cases vary in complexity and are not easily quantified, even according to elaborate point systems. Even in the Social Security adjudicatory system, where cases may be more similar to each other than those in, for example, a central panel, repeated attempts to discipline and/or remove ALJs based on low productivity statistics were overturned by the Merit Systems Protection Board, because the “SSA’s comparative statistics did not take into sufficient account the differences among these types of cases.” See Lubbers, *supra* note 6, at 599, and cases cited therein.



evaluation, and, following the logic of Prendergast's study on bias in performance evaluation,<sup>117</sup> many employees will avoid doing even this, in order to avoid potentially more negative future evaluations.

The primary purpose of any evaluation that is done should be meaningful and appropriate evaluation for self-improvement, which is preferably confidential. Any ongoing evaluation that is done should scrupulously avoid any inappropriate agency influence, and any ranking or grading of ALJs should likewise be avoided or minimized, especially if this could have any possible future effect on job security or benefits. Where merit pay for performance has been mandated by statute, especially careful attention should be paid to the standards and procedures for evaluation to assure that they avoid as much as possible any potential for inappropriate influence.

Those who do the evaluating, in any kind of system, should be trained in the importance of minimizing, and how best to minimize, the negative consequences described herein. Everyone involved in the evaluation process should acknowledge openly the very real dangers of subtle influence through evaluation, and commit to minimizing inappropriate influence through emphasis on the importance of maintaining judicial independence through strict adherence to the code of judicial conduct.

Encouraging an atmosphere of collegiality can do much to mitigate the potential negative effects of traditional performance evaluation. Allen C. Hoberg, ALJ and Director of the North Dakota Office of Administrative Hearings, North Dakota's central panel of ALJs, encourages such collegiality. Judge Hoberg does performance evaluations of ALJs, but does no ranking and indicates only satisfactory or unsatisfactory performance in various categories. Any indication of unsatisfactory performance must include specific comments, include a plan for correction, and allow for employee response. Anonymous surveys of counsel are used, but the dangers involved with them is recognized. Peer sharing, observation and feedback, with cooperation and development of professional collegiality, are also used, as are such means as staff meetings that include support staff (who also give feedback), retreats, and formal and informal training sessions. Such

---

<sup>117</sup> See *supra*, text accompanying note 71.

informal means are found to be more useful than "objective" performance evaluations.<sup>118</sup>

Others, including Chief Judge Eileen S. Willett of the Industrial Commission of Arizona, agree that ranking or grading of ALJs is unnecessary and "subject to misinterpretation or potential misuse in the political arena," and that evaluation should be aimed at self-improvement. Judge Willett suggests that "[m]icromanagement of the judicial process even in its administrative law format is a bad idea. If your selection criteria is high and the salary is commensurate with experience, your ALJ work product will usually be good. Codes of conduct, a disciplinary process that has due process and a continuing education program are also key to success."<sup>119</sup>

Deputy Director Jeff S. Masin of the New Jersey Office of Administrative Law, likewise agrees that "[a]ny attempt to rank judges, or otherwise to perform some sort of 'fine-tuned' evaluation, is neither necessary nor likely to be valid. Attempts to tie compensation to evaluation, such as providing different rates of pay increase to judges based upon some scale of performance, can cause serious internal strife and are not likely to produce valid results." A previous system of providing for different salaries arising from the evaluation process in New Jersey was changed, to eliminate such salary differences.<sup>120</sup>

In addition to the informal methods used, for example, by Judge Hoberg, it may be useful to borrow from judicial branch evaluation systems, to minimize insofar as possible the negative effects of the unilateral and inappropriately subjective aspects of traditional management-by-objectives performance evaluation, and to provide for meaningful input into, and response to, evaluations by ALJs. According to the National Center for State Courts' *State Court Organization 1993*, the evaluation methods used by the states with judicial evaluation programs include questionnaires given to lawyers, jurors, law enforcement and probation officers, social service workers,

---

<sup>118</sup>Letter from Allen C. Hoberg to Ann M. Young, (Sept. 25, 1996) (copy on file with the NAALJ Journal, at Loyola University Chicago School of Law).

<sup>119</sup>Letter from Eileen S. Willett to Ann M. Young (Dec. 9, 1996) (copy on file with the NAALJ Journal at Loyola University Chicago School of Law).

<sup>120</sup>Letter from Jeff S. Masin to Ann M. Young (Oct. 31, 1996) (copy on file with the NAALJ Journal at Loyola University Chicago School of Law).

volunteers and others who provide information to courts, litigants, staff, colleagues, appellate judges (for evaluation of lower court judges) and trial court judges (for evaluation of appellate judges); interviews with evaluated judges and others; courtroom observation; background investigations; public hearings; assessment of sentencing practices; analysis of case management data; and video tape records.<sup>121</sup>

Although some of these approaches may not be relevant in an administrative adjudication context, and may as well be cost-prohibitive in offices of ALJs that are generally much smaller and less well-funded than court systems, some might be applied appropriately, provided such potential weaknesses as those identified by Judge Stuart (in section III-B above), are acknowledged and minimized insofar as possible.<sup>122</sup> For example, consideration might be given to including the impressions of more persons than just one or two supervisors, to avoid inappropriate subjectivity as much as possible.<sup>123</sup> However, the dangers of such

---

<sup>121</sup>See *supra* note 3.

<sup>122</sup>Terms like “insofar as possible,” “as much as possible,” etc., are used in discussing potential problems and pitfalls in evaluation systems, not to indicate in any way that the utmost efforts to avoid such problems should not be undertaken, but rather to acknowledge that perfect avoidance of such problems will probably never be possible, and to note the value of such acknowledgment. If judicial arrogance is to be avoided, and counteracted through judicial evaluation, then the proponents and implementers of judicial evaluation programs should likewise be aware of the danger of becoming arrogant and oversure of the benefits of such programs.

<sup>123</sup>Rosenblum suggests the creation of evaluation panels, consisting of representatives “from bench, bar, peer groups, and relevant sectors of academia.” Rosenblum, *supra*, note 1, at 88. Overseeing such panels would be an advisory committee “drawn from such realms as retired ALJs and agency officials, current senior ALJs, federal and state court judges, private practitioners, and professors of law and of administration noted for knowledgeability about the administrative process.” *Id.* Questionnaires would be sent to parties and counsel who have appeared before evaluated ALJs, separated into pre- and post-decisional questionnaires to minimize the possibility of bias. *Id.* at 76. Panels of three members each would “consider the ALJ’s self-assessment, the questionnaire responses, [observer] reviews of the [ALJ’s] hearings, and opinions of the ALJ and the review of the employing agency,” and from all this information prepare a written evaluation “summarizing its findings about the particular strengths, particular weaknesses, recommended actions and methods for improvement, and overall quality of the ALJ’s performance.” *Id.* at 88. The report would ideally be discussed with the evaluated judge, *Id.*, and “could trigger initiation of disciplinary or removal action against an ALJ where the uniform tenor and content of evaluations present clear and convincing evidence that a particular ALJ is less than competent.” *Id.* at 89.

Rosenblum recognizes that there may be inconsistent approaches to evaluation by different evaluators: “It is not beyond likelihood, for example, that some reviewers might be favorably impressed by a judge’s meticulous concern for fairness beyond the rudiments of due

multi-rater evaluation becoming a "popularity contest" and fostering a political atmosphere,<sup>124</sup> or being subject to conscious or unconscious collusion,<sup>125</sup> must be acknowledged. Such consequences are best avoided by making programs voluntary and confidential.

To the degree any multi-rater evaluation is done, there may be differences in how different evaluators evaluate the same behavior or work product, as recognized by Professor Rosenblum,<sup>126</sup> but this should be encouraged, in the context of open and meaningful communication among evaluators where possible, and certainly between evaluators and evaluated. It should be recognized that, within certain parameters of reasonableness, there can be legitimate differences of opinion on how a judge should approach particular issues or areas of performance. After all, differing opinions are intrinsic to our legal system, as opposed to more authoritarian forms of government and law. It should be recognized that too-strong enforcement of too-rigid standards can lead

---

process requirements whereas others might view that judge's rulings as supporting and encouraging inefficient, dilatory practices. Similarly, what some might be inclined to appraise as flexibility and informality in an ALJ might strike others as inability or unwillingness to maintain adequate control over the hearing. Some might be gratified by a judge's initiative in deciding an ambiguous question of law whereas others might criticize such action as an unwarranted appropriation of an agency prerogative." *Id.* at 90.

Rosenblum recognizes that consistency in evaluation cannot be assured, but is hopeful that, through preparation of a "brief manual for evaluating judges," including consideration of input from various organizations such as the ABA National Conference of Administrative Law Judges and the Administrative Law Section of the ABA, and the Administrative Conference of the United States, along with the holding of seminars and moot evaluation proceedings, consistency could be maximized. *Id.* at 90-91.

Rosenblum's observations regarding the possibility that different evaluators might evaluate the same conduct differently are insightful. However, his answer may pose problems. Even at the level of ALJs, while certainly there are areas of consensus on appropriate judicial conduct, it is doubtful that such a universally-agreed-upon "recipe" for good judging can be arrived at, and forcing a compromise on such issues as he describes in his definition of the problem might well compromise not only independent decision-making, but also the sort of creative improvements in performance that may more effectively be produced by fostering a healthy atmosphere of collegiality in an office of judges, at whatever level. While consistent standards of evaluation are to be desired from the standpoint of avoiding bias, there may also be value in allowing for differences in style, and perhaps in this context allowing for, and even encouraging, differences of opinion among those evaluating a particular judge might make for less bias of the sort described by Prendergast.

<sup>124</sup> See *supra*, text accompanying note 89.

<sup>125</sup> See *supra* note 95.

<sup>126</sup> See *supra* note 123.

to the sort of “yes” men and women described by Professor Prendergast.

Surveys of counsel, litigants, and other participants in the adjudicatory process may provide useful feedback on perceptions of judicial performance by those who interact most closely with judges, outside their own offices. Although such surveys are perceived by many to be unreliable,<sup>127</sup> they may be helpful in addressing and avoiding a myopic perspective on judicial performance. However, great care must be taken with such surveys, to avoid any possibility of inappropriate influence on judicial performance, or compromising of judicial integrity and independence. As noted by Chief Judge Edwin L. Felter, Jr., of the Colorado Division of Administrative Hearings, such surveys should be used for developmental, self-improvement purposes, and should be confidential with regard to individual judges.<sup>128</sup>

Precautions should be taken to avoid any possible corrupting effect on independent decision-making as a result of potentially identifiable sources of positive or negative evaluations. Such surveys should not carry any consequences with regard to pay or job security, unless through such means conduct is discovered that would warrant discipline, in which event proper procedures should be followed. It should be recognized and acknowledged among all ALJs, and in any voluntary publication of the results of such surveys, that such surveys carry the inherent danger of including inappropriately negative comments from parties who do not prevail in their cases. To the degree possible, such surveys should be structured to minimize such an effect, but the potential for such results should be openly acknowledged, to prevent, as much as possible, any chilling effect on judicial independence and appropriate performance of judicial duties.

As suggested by Judge Hoberg, appropriate means of accomplishing meaningful and appropriate evaluation for self-

---

<sup>127</sup>See, e.g., the preface to the JUDICIAL PERFORMANCE EVALUATION HANDBOOK, *supra*, note 3, at iii, in which Carl O. Bradford, Chair of the National Conference of State Trial Judges, refers to the “often meaningless clamor of a bar poll.” See also Aynes, *supra* note 15, at 270.

<sup>128</sup>Edwin L. Felter, Jr., *Maintaining the Balance Between Judicial Independence and Accountability in Administrative Law*, 36 JUDGES’ JOURNAL, at 22, 54 (Winter, 1997), revised and reprinted in this issue of the Journal of the NAALJ.

improvement may include appropriate peer sharing, observation and feedback, and constructive informal counseling. However, as with surveys, peer sharing should not carry consequences with regard to pay or job security. Informality and a respectful, nonthreatening,<sup>129</sup> collegial atmosphere should be created in order to enhance the value of any peer observation and feedback that is done. If true self-improvement of judges is desired, it may be more useful to approach peer sharing as a voluntary, confidential process, rather than mandating it,<sup>130</sup> which can create interactional problems in an office. In addition, self-evaluation exercises, done with instruments created cooperatively for this purpose, may prove helpful for individual judges.

To the degree evaluations are done by chief ALJs and/or other supervising ALJs, there should also be a process for evaluation of supervisors by supervisees, with the same precautions and caveats discussed herein. This should assist in creating the sort of collegial atmosphere that is to be encouraged, and could also be done as part of an overall program of professional development for self-improvement.

All ALJs should be encouraged to discuss issues about which they have questions, recognizing that in administrative law, many unique and new issues arise that do not lend themselves to quick and easy answers. While ALJ offices should maintain a level of dignity that befits a judicial office, anything that smacks of judicial arrogance should be discouraged; it should be recognized that humility is a good quality in a judge, and is not the equivalent of weakness. Differences of opinion should be acknowledged, accepted, encouraged, and discussed openly, and all ALJs and staff in an ALJ office should be assured that no questions are "stupid," nor will any questions or

---

<sup>129</sup>It has been observed that "[j]udges are more likely to change their ways when a problem is brought to their attention in a nonthreatening way." Collins T. Fitzpatrick, *Misconduct and Disability of Federal Judges: The Unreported Informal Responses*. 71 JUDICATURE 282, 283 (1988), cited in Timony, *supra* note 3, at 655.

<sup>130</sup>It is noted that in New Jersey, a committee on the peer review system of evaluating judicial writing recommended against peer review, suggesting that only with one reader could there be some assurance of the use of consistent standards of evaluation. However, the committee stated that the system at the time was "doomed to failure so long as it is in any way tied to financial reward or reappointment decisions." Peer sharing with no effect on job security or benefits, undertaken in a non-threatening atmosphere of collegiality, should not suffer from the same problem. Report of the Committee on Writing Evaluation, at 10. Copy attached to October 31, 1996, letter from Jeff Masin. See *supra*, note 120.

disagreement on issues be held against them later<sup>131</sup> in any performance evaluation or other context. This is necessary, to encourage more effective performance of all duties at all levels.

Finally, constructive informal counseling and mentoring, along with training, educational and professional development activities in areas of perceived weakness, may be used with ALJs who have problems with performance, just as informal mentoring, observation and feedback can be used with new ALJs, to assure better performance. This should be done as much as possible in the same non-threatening, collegial atmosphere that is described above, in order to achieve the best results.

### E. Complaint and Disciplinary Procedures

Any problems deemed serious enough to affect an ALJ's job should be handled through fair complaint and disciplinary procedures. A system for providing an easy, but formalized, means of making complaints against ALJs who are perceived as violating standards of conduct or otherwise engaging in inappropriate conduct, and of appropriately evaluating and resolving complaints against ALJs, should be developed, which does not minimize serious complaints, but also does not overemphasize complaints relating to the merits of cases, or otherwise compromise independent decision-making.<sup>132</sup> All complaints should be required to be written, to minimize problems of hearsay and interpretation, and to mitigate any problem of frivolous complaints.

Any discipline and removal of ALJs should be accomplished according to clearly defined procedures, based on the same principles of due process and independent decision-making that exemplify the American legal system and that ALJs should apply in the cases they hear, and should be based on standards of conduct set forth in a formally-adopted code of judicial conduct. Principles applicable to the discipline of judges in the judicial system may and should be applied to ALJs, although particular procedures may differ. For example, the Model Act Creating a State Central Hearing Agency, adopted by the

---

<sup>131</sup>As Kaufman has observed in the context of judicial discipline, "there would be an inevitable loss of frankness [between judges] if each participant feared that candor might one day build a case against him." Kaufman, *supra* note 37, at 711.

<sup>132</sup> See Felter, *supra* note 128, at 23.

American Bar Association House of Delegates at the mid-year meeting in San Antonio on February 3, 1997, provides for "notice and an opportunity to be heard in an Administrative Procedure Act or other statutory-type hearing and a finding of good cause by an impartial hearing officer," prior to any action to remove, suspend, demote, or take any disciplinary or other adverse action against an ALJ, including any action that might later influence a reduction in force.<sup>133</sup>

A commission, panel, or "court of the administrative judiciary" might also be established, either as an alternative or an adjunct to any other procedures used, to consider complaints against ALJs, and issue ethical opinions. ALJ offices might also consider approaching such entities in the judicial branch, to establish working relationships, on whatever level deemed appropriate in the jurisdiction.

The use of evaluations in connection with any discipline of ALJs should be avoided, in order to avoid the negative consequences discussed above, and to maximize the developmental, self-improvement value of any evaluations that are done. Any evaluation sought to be used in any disciplinary procedure against any ALJ should be considered hearsay unless a specific evidentiary exception to hearsay applies.

In sum, if a complaint involves anything other than frivolous allegations, it deserves serious attention, both in the interest of the public and the judge involved, and should result in discipline or removal only after notice and an opportunity for a hearing, with competent proof of wrongdoing or incompetence, and a determination by an independent decision-maker or decision makers, applying the principles of the American legal system.

## V. Conclusion

The American legal system is designed to treat similarly-situated persons as similarly as possible. This goal addresses a fundamental human concern: that of fairness, a concern so basic that young children understand it implicitly and sometimes demand it insistently. Present calls for accountability and evaluation of judges, from a public that is

---

<sup>133</sup> See *supra* note 107.



discontent with our legal system, may indeed perhaps best be understood as expressions -- more or less articulate -- of frustration with a system that is perceived to be unfair; the most obvious example of this is the very understandable concern for victims of crime. The arguments of many in the academic and governmental communities who advocate evaluation of ALJs are similarly often rooted in concerns of fairness.<sup>134</sup> The administrative legal system, like the judicial branch, must address these frustrations and concerns.

Our legal system includes numerous procedures and requirements to assure that decisions are made fairly, the most important of which is the participation of an independent decision-maker. The specifics of other procedural protections of various constitutional rights may, and no doubt will, be forever debated, as will the proper definition and particulars of judicial independence. However, without judicial independence, the basic fairness the public cries out for cannot be achieved. And if the public, confronted with a multitude of painful, sometimes complex problems, does not see the value of judicial independence clearly, it is our responsibility as administrative law judges and members of the legal/judicial community, to clarify the issue. Without such clarification, even the best judicial evaluation system will likely do little to assuage public concern about the judiciary, including the administrative judiciary.

Instead, inappropriate evaluation of judges may actually lead to greater unfairness, by compromising independent decision-making. Any system of standards setting and enforcement for judges, including ALJs, should be approached with this danger clearly in mind, in a very real and practical sense.

Where evaluation of ALJs is mandated, close and careful attention must be paid to the criteria and methods of evaluation that are used, to avoid compromising decisional independence as much as possible. Most importantly, whatever criteria or methods are used, ranking and grading of ALJs should be avoided or minimized as much as possible, as should tying evaluations to job security or benefits, in order not to undermine independent decision-making, productivity, efficiency, collegiality, and quality in work product. The use of surveys and other

---

<sup>134</sup>See *supra* note 7.

multi-rater systems may alleviate the potential for bias that may inhere in evaluation by only one or two supervisors, but the use of multi-rater systems does not eliminate any possibility of bias or inappropriate influence, and can cause additional problems, which should be acknowledged and addressed. Confidential, developmental evaluation for self-improvement, including self-evaluation, peer observation and sharing, informal counseling and feedback, along with training and professional development activities in areas of perceived weakness, are better approaches for improving judicial performance.

Where it is possible to avoid or minimize the use of traditional management-by-objectives performance evaluation models, well-thought-out alternative systems, developed collegially, which provide for rigorous selection of ALJs, high standards of conduct, effective continuing education and professional development, appropriate and meaningful observation and feedback, fair evaluation of complaints, and appropriate discipline and removal procedures, should more than adequately accomplish the purposes of traditional performance evaluation systems, without resulting in the negative consequences to independent decision-making that can be entailed in purely traditional models.

Indeed, what we can learn from current management theory and research suggests that ALJs who meet rigorous selection standards, maintain high standards of conduct and professionalism through such means as continuing education programs and peer sharing, are aware of fair and appropriate disciplinary and removal procedures, but are not subject to traditional periodic performance evaluations, should be at least as, if not more motivated to perform well as they would be if they were subject to traditional management-by-objectives performance evaluation, and should be more motivated to make all rulings and decisions independently, according to the values of due process, based solely on the facts and law of the cases before them. An alternative system of selection, standards-setting and enforcement may, if appropriately implemented, actually strengthen "weak" ALJs, by encouraging collegial sharing and cooperation, enhancing the responsibility of independent decision-making, and fostering professional development and pride in work product.

Creating standards-setting and enforcement systems that take into

account the basic principles of the American legal system and what we can learn from the field of management should thus more effectively both improve ALJ performance and encourage ALJ fulfillment of the responsibility of independent, impartial, and fair decision-making. This approach, along with appropriate public education activities, should thereby better and more appropriately address public discontent with the legal system, through reassurance that the system is in fact fair, as it is intended to be.

If effective, independent decision-making is to be desired in administrative adjudication, it should not be hampered by unintended negative consequences of traditional evaluation systems, but rather should be encouraged through a system of appropriate selection, standards-setting and enforcement procedures, which will also better assure efficiency and excellence in performance. If independent decision-making is not actively and systematically fostered, what we offer the public is the deception that decisions are independent when in fact they may not be. The pejorative expression, "good enough for government service," is not an appropriate maxim when due process, equal justice under the law, and independent decision-making are at issue. As administrative law judges, we should accept the responsibility of independent decision-making, and the challenge of creating systems that foster the effective fulfillment of this significant and far-reaching responsibility.