Chapter 13

HOW AND WHY LABOR ARBITRATORS DECIDE DISCIPLINE AND DISCHARGE CASES

I. An Empirical Examination

Laura J. Cooper, Mario F. Bognanno, and Stephen F. Befort*

Introduction

Often without attribution, the arbitration literature includes statements about how "most," "many" or "some" arbitrators decide discipline and discharge cases. *Elkouri & Elkouri* is probably the best known source of such generalizations.[1] Although statements like these are often consistent with conventional wisdom, empirical support is seldom provided. To illustrate, consider the first of three examples taken from the National Academy of Arbitrators' *The Common Law of the Workplace*:

> Because industrial discipline is corrective rather than punitive, *most* arbitrators require use of progressive discipline, even when the collective agreement or employment contract is silent on the subject.[2]

Based on personal experiences, this statement rings true. But is it? Our new study of 2,055 discipline and discharge cases includes 454 decisions in which the arbitrator found just cause for a lesser discipline due to the presence of one or more mitigating factors. Using these data, Table 1 reports the results of two different tests

*Listed in reverse alphabetical order, the authors are on the faculty of the University of Minnesota, Minneapolis, Minnesota. Ms. Cooper is the J. Stewart and Mario Thomas McClendon professor, Law School; Mr. Bognanno is professor emeritus, Industrial Relations Center, Carlson School of Management; and Mr. Befort is the Gray, Plant, Mooty, Mooty, and Bennett professor, Law School. E-mails: <lcooper@umn.edu>; <bogna001@umn.edu>; and <befor001@umn.edu>. The authors would like to thank the staff of the Minnesota Bureau of Mediation Services and, particularly, Commissioner James A. Cunningham and Assistant to the Commissioner Carol S. Clifford, for their assistance and facilitation of access to the state's arbitration case files. We also wish to thank Jonathan Booth, Deborah Broderson, Tana Dietzler, Jessica Marsh, Kristen Munday, Thomas Norman and Jeff Vockrodt for their research assistance.

[1]Ruben, ed., Elkouri & Elkouri: How Arbitration Works, 6th ed (BNA Books 2003).

[2]St. Antoine, ed., The Common Law of the Workplace: The Views of Arbitrators, 2d ed. (BNA Books 2005), at §6.7 Comment (emphasis added).

of the referenced progressive discipline generalization in the *Common Law of the Workplace*. The first test involved computing the percent of the 454 decisions that mention a listed mitigating factor. (See Table 1, column 3.) In the second test, we computed the percent of all mitigations cited (810) that note a listed mitigating factor. (See Table 1, column 4.) As both tests show, the "lack of progressive discipline" was a major mitigating factor that lead to disciplinary reductions in 27.09 percent of all cases and in 15.19 percent of all mitigations cited. According to both tests, the dominant reasons arbitrators gave for reducing discipline were "punishment too severe for offense" (33.92 percent of all cases and 19.01 percent of mitigations cited) and "good work record" (32.60 percent of all cases and 18.27 percent of mitigations cited). Regardless of the test-measure used, the rank ordering of the relative importance of each mitigating factor remains the same.

Table 1: Mitigating Factors for Reducing Employer Discipline

Mitigating Factors	N*	Rank Order	% of Decisions	% of Mitigations Cited*
Punishment too severe for offense	154	1	33.92	19.01
Good work record	148	2	32.60	18.27
Lack of progressive discipline	123	3	27.09	15.19
Length of service	106	4	23.35	13.09
Other	84	5	18.50	10.37
Absence of serious harm caused by employee's conduct	65	6	14.32	8.02
Employee acted in good faith	59	7	13.00	7.28
Rehabilitation (i.e., drug, alcohol, mental health, and gambling addiction)	49	8	10.79	6.05
Provocation	22	9	4.85	2.72
Total	810	NA	NA	100.00

*Σ N = 810 as more than one mitigating factor is cited in some of the 454 case decisions.

A second example from the *Common Law of the Workplace* is: "*Generally*, arbitrators will enforce last-chance agreements."[3] We can directly test the validity of this assertion. Our study coded answers to the questions, "What was the outcome of the award?" and "Was the employee, at the time of the discipline, working under a 'last chance' condition?"[4] As Table 2 shows, 165 employees in our study were working under the terms of a last-chance agreement at the time of their discipline and the balance were not. The share of "management win" outcomes approximated 50 percent for disciplinary cases in general and also for employees who were not in a last-chance situation. However, when arbitration occurred in the presence of a last-chance agreement, the percentage of "management wins" increased more than 20 percentage points to 70.91 percent, while "union wins" and "split decisions" each fell by about 10 points. These results provide compelling statistical support for the *Common Law of the Workplace* assertion that arbitrators "will enforce last-chance agreements."[5]

Finally, there is the quote:

> Once it is determined that a back pay award is appropriate, an arbitrator may remand the task of computation to the parties. Such a remedy is *usually*, but not always, accompanied by retention of jurisdiction by the arbitrator in the event that there is a subsequent dispute over the amount.[6]

Assuming that arbitrators generally remand the task of computing back pay to the parties, we can test this proposition directly. Our sample of arbitration decisions includes 432 cases in which the arbitrator reinstated the employee with either full or partial back pay. With this subset of cases in mind, we coded whether "the arbitrator's award explicitly retains jurisdiction to decide subsequent disputes . . . with regard to the remedy." We were somewhat surprised to learn that in three out of four (73.38 percent) of these cases, the arbitrator did *not* retain jurisdiction, suggesting

---

[3] *Id.* at §6.29 Comment (emphasis added).

[4] An arbitration "outcome" refers to a case's result, whether management won, union won, or the arbitrator issued a split decision, implying that the arbitrator found just cause for lesser discipline than that imposed by the employer.

[5] We use the chi-square test to make comparisons between two attributes to determine whether there is any relationship between them. For example, Table 2 involves the comparison between outcomes and whether the employee's continuing employment was subject to the terms of a last-chance agreement. Implicitly, we are interested in knowing whether the presence of a last-chance agreement affects the arbitration outcome. With a chi-square of 32.5155 and 2 degrees of freedom, the chances are less than 1 in 1,000 (i.e., $p = 0.000$) that there is no relationship between outcomes and last-chance agreements.

[6] Common Law of the Workplace, *supra* note 2, at §10.14 (emphasis added).

Table 2: Arbitrator's Award in Discharge and Discipline Cases by Last-Chance Agreement*

(N-size, row %, and column % are the top, middle, and bottom value, respectively.)

Arbitration Outcome	Last-Chance	No Last-Chance	Row Totals
Arbitrator Found Just Cause for Discipline (Management Win)	117 11.45 70.91	905 88.55 47.88	1,022 100.00 49.73
Arbitrator Found No Just Cause for Any Discipline (Union Win)	18 4.08 10.91	423 95.92 22.38	441 100.00 21.46
Arbitrator Found Just Cause for Lesser Discipline than that Imposed by Employer (Split Decision)	30 5.07 18.18	562 94.93 29.74	592 100.00 28.81
Column Totals	165 8.03 100.00	1,890 91.97 100.00	2,055 100.00 100.00

*chi-square (2) = 32.5155, p = 0.000

that the quoted assertion may be questioned and is properly the subject for future research.

The *Common Law of the Workplace* is replete with affirmative claims like these. Indeed, elsewhere in this paper we put more of them to a test of empirical validation. The point, however, is that these assertions led us to identify the need for a large, wide-ranging survey of the content of discipline and discharge decisions, which is rare in the arbitration literature. In mid-2004, the Academy's Research and Education Foundation awarded a grant to support this study and, for the next two years, we coded an enormous amount of information in 2,055 discipline and discharge decisions, as well as information about the arbitrators who decided them. The information we used to critique the above quotes from the *Common Law of the Workplace* are from this survey. We are now analyzing these data. The purpose of this paper is to present an introductory glimpse at its rich breadth, depth, and analytical potential.

In the next section of this paper, we discuss the survey, briefly present some summary statistics that characterize the data, and, in light of relevant literature, we put our sample to work as the judge in two of the literature's "win, lose or split" debates. In addition, we present new tabulations on the relationship between discharge case remedies and whether (1) the case arose in the public or private sector, (2) the arbitrator had a legal education, and (3) the arbitrator was a NAA member at the time of the decision. In the third section of this paper, we take a critical look at compromise or split-decision remedies in discharge cases. The fourth section of this paper addresses arbitrator and advocate use of the Seven Tests and arbitrators' application of various standards for quantum of proof. We conclude with summary remarks in the final section.

Methods and Descriptive Statistics

Survey Sample. As previously observed, we have coded a new and unusually large sample of 2,055 discipline and discharge decisions issued between 1982 and 2005 by 81 different arbitrators. All of these decisions involved Minnesota public and private sector bargaining units. Since the early 1980s, arbitrators on Minnesota's Bureau of Mediation Services (BMS) roster have been required to file a copy of their decisions with the BMS regardless of the source of the arbitrator's appointment. Thus, our discipline and discharge sample is a subset of the several thousand decisions of all types that are archived in BMS case files, and generally it approximates the population of *all* discipline and discharge decisions issued in Minnesota during the relevant time period.[7]

Since late in 1982, the BMS or its vendor, LRP Publications, has prepared written summaries of each arbitration decision, cataloging each in loose-leaf binders entitled *Minnesota Labor Arbitration Awards Summaries.* Each summary is given an identification number, which codes key aspects of the decision, including the date and sector of the case, and whether the decision favors management, labor, or is a split decision. Further, a "Key Classification Index" (KCI) number is assigned to each decision. Our sample consists

[7]Minnesota Rules, Chapter 5530.08, subpart. 9, requires arbitrators to file copies of their public and private sector awards with the Commissioner of the BMS. Specifically, the rule requires: "Unless one or both private sector parties have specifically requested that an award not be provided to the commissioner, arbitrators shall submit copies of all awards involving Minnesota work sites to the commissioner regardless of the source of appointment or selection. Awards filed with the commissioner are public documents." We presume that some arbitrators on the BMS roster did not comply fully with this Rule.

of all 1982–2005 decisions assigned KCI numbers 37 (discipline), 109 (suspension), and 112 (termination of employment).

Our research team included four second- and third-year law students who retrieved from BMS files a hard copy of the KCI-designated cases. Trained to analyze and code, the legal assistants coded 36 survey items for each decision, to which we later merged information about the arbitrator's age, education, gender, occupation, and Academy membership status.[8] Due to coding complexities, we omitted arbitration cases that involved multiple grievants. While going through the BMS case files, our legal assistants unearthed a set of veterans' termination and teacher tenure cases governed by "just cause" standards that had not been indexed. These awards were also coded for analysis and are included in our sample of 2,055 awards. In all, we collected data on more than 100 variables per award.[9]

Our sample has some special qualities. Our discipline and discharge decisions are not limited to cases published by the Bureau of National Affairs (BNA) and Commerce Clearing House, Inc. (CCH). Published decisions are not representative of unpublished decisions, which are a super-majority of all decisions issued. In a study uniquely designed to test this proposition, Jack Stieber and colleagues found that discharge decisions published by the BNA and CCH were not representative of unpublished Michigan discharge decisions, at least not during 1979–1982.[10] Yet most field studies of discipline and discharge arbitration are based on published BNA and CCH decisions. Some parties permit the publication of all of their decisions, while others occasionally or never do.

---

[8]To ensure consistent coding, we did some cross-checking. We three and our legal assistants read several specimen decisions, coding each. We then matched our codes on an item-by-item basis, checking for differences. Where they existed, we agreed on a coding rule and, on occasion, we also honed the phrasing of survey items.

[9]For example, with respect to the "offense" variable, our survey allowed for the coding of 43 different categories. In this paper, we clustered similar offenses into the following nine categories: Violence & Aggression (467 cases), Attendance & Absenteeism (352 cases), Dishonesty (377 cases), Drug & Alcohol Use (96 cases), On-the-Job Misconduct (175 cases), Off-the-Job Misconduct (94 cases), Insubordination (469 cases), Performance Issues (370 cases), and Other (24 cases).

[10]Stieber, Block & Corbitt, *How Representative are Published Decisions?*, in Arbitration 1984: Absenteeism, Recent Law, Panels, and Published Decisions: Proceedings of the 37th Annual Meeting, National Academy of Arbitrators, ed. Gershenfeld (BNA Books 1985), at 172. Regarding Michigan's discharge cases, these researchers found, for example, that published decisions under-represent public sector decisions and female grievants, and over-represent cases in which unions use attorneys and in which briefs are filed. In addition, they found that neither BNA- nor CCH-published discharge decisions were representative of unpublished Michigan awards. To illustrate, the "management win" rate of BNA and CCH discharge decisions were significantly higher and lower, respectively, than the "management win" rate for unpublished Michigan discharge cases.

Moreover, with the concurrence of the parties, some arbitrators submit some or all of their awards for publication, while others never do. Finally, of course, publishers control whether to publish decisions they receive, and their selection criteria are not designed to ensure that published cases are representative.[11]

Moreover, our sample approximates the population of all discipline and discharge decisions rendered in Minnesota, or nearly so. Thus, representation at a Minnesota-level of generality is not an issue, except that public sector decisions in our sample are probably over-represented.[12] However, as each state's industrial relations and arbitrator characteristics are uniquely different and change in different ways over time, our sample of discipline and discharge cases is not necessarily representative of national-level decisions and the arbitrators who decided them. Sampling bias like this limits our ability to make reliable national inferences because we do not know the likely effects of this bias as, to the best of our knowledge, a nationwide sample of unpublished discipline and discharge decisions and the characteristics of the arbitrators who decided them has never been taken. Nevertheless, because our sample is one of the largest and most detailed compilations of published and unpublished discipline and discharge decisions ever assembled, it should serve as a reasonable basis for testing assertions like those discussed earlier. In addition, it is used to re-test some of the literature's more interesting hypotheses, most of which were based on relatively small samples of published decisions.

Sample Characteristics. We presented the relative frequency distributions of selected sample characteristics in Tables 1 and 2. Additional summary characteristics are presented throughout the balance of this paper. At this point, we list several other characteristics of interest. The 2,055 decisions in our data set were issued over a 24-year period by 81 different arbitrators. Select characteristics of these 81 arbitrators include the following:

---

[11]Bernstein, *Proper Prepublication Procedures: An Arbitrator's Comment*, in Arbitration 1984: Absenteeism, Recent Law, Panels, and Published Decisions: Proceedings of the 37th Annual Meeting of the National Academy of Arbitrators, ed. Gershenfeld (BNA Books 1985), at 192.

[12]Some private sector parties may have sometimes exercised their right not to permit BMS submission. Arbitrators not on the BMS roster deciding cases arising from Minnesota work sites are under no obligation to send decisions to the BMS. We believe that arbitrators not on the BMS roster are primarily deciding private sector cases.

- 17.28 percent were women who decided 17.03 percent of the cases.
- The mean age of arbitrators at the time of the award was 57.21 years.
- 65.43 percent had law degrees and decided 61.75 percent of the cases.
- 8.60 percent had Ph.D. degrees and decided 9.05 percent of the cases.
- 27.16 percent were academics who decided 40.05 percent of the cases.
- 32.09 percent were NAA members who decided 70.61 percent of the cases.

Finally, in 68.95 percent of the cases arbitrators were selected through three appointing agencies: BMS, Federal Mediation and Conciliation Service (FMCS), and American Arbitration Association (AAA). Veterans' preference and teacher tenure cases made up 4.52 percent of all decisions.

Win, Lose, or Split in the Empirical Literature. Table 3 presents our "management win," "union win," and "split decision" results, which, for different reasons, are of interest to the parties, professional arbitrators, and academics. For comparison purposes we also present the results reported in three contemporary studies. Using as a guide the latest publication date of the awards analyzed in each of the contemporary studies, it appears as though the rate of "union win" outcomes reached a maximum around 1995 and, as suggested by our study, has since been trending downward; whereas, the "management win" and "split decision" outcomes have been creeping upward. But appearances can be deceiving.

This conclusion can be tested with precision because our data cover the years from 1982 to 2005. Accordingly, we partitioned our data into three 8-year time periods—1982–1989, 1990–1997, and 1998–2005—and then analyzed arbitration outcomes by time period and, subsequently, by time period and sector of the case (public or private). Table 4 shows that during the three time periods under consideration, the inter-period "management win" rate jumped upward quite dramatically between 1982–1989 (45.70 percent) and 1990–1997 (51.63 percent), but thereafter it has remained relatively flat (50.85 percent). Further, over time the "union win" rate has remained relatively stable, with the inter-period increase in the share of "management win" awards resulting from a decline in the share of "split decision" outcomes. Overall,

TABLE 3: Discipline and Discharge Arbitration Outcomes by Study: Management Win, Union Win, or Split

(N-size and row % are the top and bottom value, respectively.)

Study	Management Win	Union Win	Split Decisions	Total
1986 Data—Zirkel and Breslin[1]	108 46.8	43 18.6	80 34.6	231 100.00
1988-1990 Data—Mesch and Dalton[2]	193 44.6	128 29.5	112 25.9	433 100.00
1993 Data—Mesch[3]	1,888 47.8	1,282 32.5	779 19.7	3,949 100.00
1982-2005 Data— Cooper, Bognanno and Befort	1,022 49.7	441 21.5	592 28.8	2,055 100.00

[1]Zirkel & Breslin, *Correlates of Grievance Arbitration Awards*, 24 J. Collective Negotiations 45 (1995) (data from designated 1986 volumes of BNA, Labor Arbitration Reports, and CCH, Labor Arbitration Awards).

[2]Mesch & Dalton, *Arbitration in Practice: Win, Lose, or Draw?*, 4 The Human Resources Prof. 37 (1992) (data from BNA, Labor Arbitration Reports, 1988–1990).

[3]Mesch, *Grievance Arbitration in the Public Sector: A Conceptual Framework and Empirical Analysis of Public and Private Sector Arbitration Cases*, 15 Rev. Pub. Pers. Admin. 22 (1995) (data from BNA, Labor Arbitration Reports, 1987–1993).

it seems that the "union win" rate has been fairly stable, holding at around 21.70 percent for the past 15 years or so. Note, however, in this case there is a 7.1 percent chance that the identified time periods and outcomes are independent (i.e., not related).

Next, we compare our results with those in Debra Mesch's 1995 study, which examined differences between private and public sector discipline and discharge arbitration outcomes. Table 5 presents her overall results, which show a significant difference in outcomes by sector.[13]

[13]Mesch, *Grievance Arbitration in the Public Sector: A Conceptual Framework and Empirical Analysis of Public and Private Sector Arbitration Cases,* 15 Rev. Pub. Pers. Admin. 23, 30 (1995).

Table 4: Discipline and Discharge Arbitration Outcomes by Time Periods: Management Win, Union Win, or Split*

(N-size, row %, and column % are the top, middle, and bottom value, respectively.)

Time Periods	Management Win	Union Win	Split Decisions	Row Totals
1982–1989	266 45.70 26.03	123 20.96 27.66	194 33.33 32.77	582 100.00 28.32
1990–1997	458 51.63 44.81	192 21.65 43.54	237 26.72 40.03	887 100.00 43.16
1998–2005	298 50.85 29.16	127 21.67 28.80	161 27.47 27.20	586 100.00 28.52
Column Totals	1,022 49.73 100.00	441 21.46 100.00	592 28.81 100.00	2,055 100.00 100.00

*chi-square (4) = 8.6283, p = 0.071

Table 5: Mesch's Outcomes by Sector for Disciplinary Cases*

Arbitral Outcome	Private Sector N %		Public Sector N%		Totals N%	
Management Win	1,390	49.3	498	44.2	1,888	47.8
Union Win	867	30.7	415	36.8	1,282	32.5
Split Decision	565	20.0	214	19.0	779	19.7
Totals	2,822	71.5	1,127	28.5	3,949	100.00

*chi-square (2) = 13. 99, p < .001

These data are striking in that union win rates are far higher in the public sector than in the private sector—36.8 percent versus 30.7 percent. However, this difference is inconsistent with the findings of several other researchers, all of whom used multivariate analytical methods.[14] However, to be fair, these findings were premised on *discharge* data, not *discipline and discharge* data.[15] Accordingly, we thought it would be interesting to contrast the results in Table 5 with our own.

Table 6 presents contrasting results, using both our aggregate 1982–2005 data, and data for the three 8-year time periods introduced earlier. Our aggregate results differ markedly from Mesch's. Our public sector union win rate is lower, not higher, than the private sector union win rate: 20.69 percent versus 22.58 percent, however, the identified bivariate relationship is suspect, statistically speaking. But a perusal of our 1982 percent 1989 results suggest that this was not always so. During this time period, the union win rate was much higher in the public sector than the private sector: 25.26 percent versus 16.61 percent. This result is significant and it parallels Mesch's result; but as Table 6 also shows, at some point during the subsequent two time periods, the private sector union win rate surpassed the public sector union win rate. Explaining this phenomenon is a subject for future research. Rhetorically, we ask, "Have public sector unions been taking relatively weaker cases to arbitration than their private sector brethren?"

---

[14]Nelson & Uddin, *The Impact of Delay on Arbitrators' Decisions in Discharge Cases*, 3 Lab. Stud. J. 4, 15 (1998); Zirkel & Breslin, *Correlates of Grievance Arbitration Awards*, 24 J. Collective Negotiations 45, 50 (1995) ; Bemmels, *The Effect of Grievants' Gender on Arbitrators' Decisions*, 41 Indus. & Lab. Rel. Rev. 251, 257 (1988); and Block & Stieber, *The Impact of Attorney and Arbitrators on Arbitration Award*, 40 Indus. & Lab. Rel. Rev. 543, 549 (1987).

[15]This distinction is relevant because in our data, union win rates were significantly higher in suspension and reprimand cases than in discharge cases.

Table 6: Discipline and Discharge Arbitration Outcomes by Sector and Time Periods: Win, Lose, or Draw

(N-size and row % are the top and bottom value, respectively.)

Awards	Management Win	Union Win	Split Decisions	Row Totals
1982–2005[1]	403	189	245	837
Private Sector	48.15	22.58	29.27	100.00
Public Sector	619	252	347	1,218
	50.82	20.69	28.49	100.00
1982–1989[2]	137	48	104	289
Private Sector	47.40	16.61	35.99	100.00
Public Sector	129	74	90	293
	44.03	25.26	30.72	100.00
1990–1997[3]	177	77	91	345
Private Sector	51.30	22.32	26.38	100.00
Public Sector	281	115	146	542
	51.85	21.22	29.94	100.00
1998–2005[4]	89	64	50	203
Private Sector	43.84	31.53	24.63	100.00
Public Sector	209	63	111	383
	54.57	16.45	28.98	100.00

[1]chi-square (2) = 1.6446, p = .439
[2]chi-square (2) = 6.7647, p = .034
[3]chi-square (2) = 0.1548, p = .927
[4]chi-square (2) = 17.8344, p = .000

Another issue considered in the previous literature concerns arbitrator neutrality and the gender of the grievant. In 1988, Brian Bemmels analyzed the effect of the grievant's gender on outcomes, using data from 104 discharge cases in Alberta, Canada, from 1981–1983. Most arbitrators would reject the notion that their decisions are influenced by personal gender prejudices. Bemmels' fully controlled analytical models, although generally

affirming Allen Ponak's 1987 findings based on 159 Alberta, Canada, discharge cases,[16] specifically establish that:

> [W]omen were twice as likely as men to have their grievances sustained and 2.7 times more likely to receive full reinstatement, and in the cases in which a suspension was imposed by the arbitrator, women received, on average, a suspension 2.1 months shorter than for men.[17]

In contrast, none of the discharge studies by Block and Stieber (1987),[18] Scott and Shadoan (1989),[19] and Nelson and Uddin (1998)[20] found a significant relationship between the grievant's gender and the arbitration outcome. The same is true for Zirkel and Breslin (1995),[21] who used discipline and discharge data.

Table 7 depicts the arbitration award and remedy by gender for the 1,432 discharge cases in our sample. It appears to the naked eye that the distribution of awards and remedies by gender are not meaningfully different. Indeed, the male and female column distributions are statistically independent. From these results it appears that the arbitrators in our sample did not decide cases on the basis of gender. However, a cautionary note is warranted simply because this absence of a gender effect may prove to be unstable once other factors are introduced into the analysis.

We conclude this section of the paper with a presentation of the relationship between an array of remedy outcomes in discharge cases and three distinct variables of interest, namely, sector (public and private); whether the arbitrator has a J.D. degree; and whether the arbitrator was a member of the NAA at the time the case was decided. The chi-square tests adjacent to Tables 8, 9, and 10 show a strong relationship between remedial outcomes and each of the named variables.

Our data include an almost equal number of discharge cases in the public (705) and private (727) sectors. Table 8 shows that generally arbitration outcomes are more favorable to public sector

---

[16]Ponak, *Discharge Arbitration and Reinstatement in the Province of Alberta*, 42 Arb. J. 39 (June 1987).

[17]Bemmels, *supra* note 14, at 251, 259.

[18]Block & Stieber, *supra* note 14, at 543, 549 (analyzing 1,213 published BNA and CCH cases, and unpublished Michigan discharge cases issued between 1979 and 1982).

[19]Scott & Shadoan, *The Effect of Gender on Arbitration Decisions*, 10 J. Lab. Res. 429, 433 (1989).

[20]Nelson & Uddin, *The Impact of Delay on Arbitrators' Decisions in Discharge Cases*, 3 Lab. Stud. J. 3, 15 (1998).

[21]Zirkel & Breslin, *Correlates of Grievance Arbitration Awards*, 24 J. Collective Negotiations 45, 50 (1995).

Table 7: Discharge Arbitration Outcomes by Gender: Management Win, Union Win, or Split*

(N-size, row %, and column % are the top, middle, and bottom value, respectively.)

Arbitration Award	Males	Females	Row Totals
Management Win: Discharge Upheld	573 76.30 52.91	178 23.70 51.00	751 100.00 52.44
Union Win: Reinstate with Full Back Pay	207 73.14 19.11	76 26.86 21.78	283 100.00 19.76
Split Decision: Reinstate with Partial Back Pay	108 72.48 9.97	41 27.52 11.75	149 100.00 10.41
Split Decision: Reinstate with No Back Pay	191 79.58 17.64	49 20.42 14.04	240 100.00 16.76
Split Decision: Entitled to a Future Vacancy	4 44.44 0.37	5 55.56 1.43	9 100.00 0.63
Column Totals	1,083 75.63 100.00	349 24.37 100.00	1,432 100.00 100.00

*chi-square (2) = 8.7145, p = .069

employers than to private sector employers. Public sector employers have a 56.17 percent win rate, while private sector employers prevail in 48.83 percent of cases.

Table 9 depicts the relationship between remedial outcomes and whether the arbitrator is trained as an attorney. Arbitrators who have earned a J.D. degree are considerably more inclined to rule in favor of unions in discharge cases than are their non-attorney counterparts. Thus, employers prevailed in almost 60 percent

TABLE 8: Remedial Outcome by Sector*

(N-size, row %, and column % are the top, middle, and bottom value, respectively.)

Reinstatement Outcomes	Public Sector	Private Sector	Total
Uphold the Discharge	396 52.73 56.17	355 47.27 48.83	751 100.00 52.44
Reinstate—Full Back Pay	119 42.05 16.88	164 57.95 22.56	283 100.00 19.76
Reinstate—Partial Back Pay	75 50.34 10.64	74 49.66 10.18	149 100.00 10.41
Reinstate—No Back Pay	109 45.42 15.46	131 54.58 18.02	240 100.00 16.76
Entitlement to Future Vacancy	6 66.67 0.85	3 33.33 0.41	9 100.00 0.63
Total	705 49.23 100.00	727 50.77 100.00	1,432 100.00 100.00

*chi-square (4) = 12.0821, p = .017

of discharge cases before non-attorney arbitrators, while achieving a similar outcome in fewer than 50 percent of such cases before attorney arbitrators.

As shown in Table 10, the impact of NAA membership on remedial outcomes is significant, although less marked in absolute terms relative to the two previously discussed variables. NAA membership is often used as a proxy for arbitral experience, as an arbitrator must have decided a minimum of 50 cases within a five-

Table 9: Remedial Outcome by Arbitrator's Attorney Status*

(N-size, row %, and column % are the top, middle, and bottom value, respectively.)

Reinstatement Outcomes	Non-Attorney Arbitrator	Attorney Arbitrator	Total
Uphold the Discharge	302 40.21 59.10	449 59.79 48.75	751 100.00 52.44
Reinstate—Full Back Pay	96 33.92 18.79	187 66.08 20.30	283 100.00 19.76
Reinstate—Partial Back Pay	38 25.50 7.44	111 74.50 12.05	149 100.00 10.41
Reinstate—No Back Pay	71 29.58 13.89	169 70.42 18.35	240 100.00 16.76
Entitlement to Future Vacancy	4 44.44 0.78	5 55.56 0.54	9 100.00 0.63
Total	511 35.68 100.00	921 64.32 100.00	1,432 100.00 100.00

*chi-square (4) = 29.5003, p = 0.00

year period in order to be eligible for NAA membership. Arbitrators who are NAA members, a group that decided approximately 54 percent of the discharge cases in our sample, ruled in favor of management slightly more often (54.15 percent) than did non-NAA members (50.45 percent).

TABLE 10: Remedial Outcome by Arbitrator's NAA Status*

(N-size, row %, and column % are the top, middle, and bottom value, respectively.)

Reinstatement Outcomes	Non-NAA Arbitrator	NAA Arbitrator	Total
Uphold the Discharge	333 44.34 50.45	418 55.66 54.15	751 100.00 52.44
Reinstate—Full Back Pay	151 53.36 22.88	132 46.64 17.10	283 100.00 19.76
Reinstate—Partial Back Pay	87 58.39 13.18	62 41.61 8.03	149 100.00 10.41
Reinstate—No Back Pay	83 34.58 12.58	157 65.42 20.34	240 100.00 16.76
Entitlement to Future Vacancy	6 67.67 0.91	3 33.33 0.39	9 100.00 0.63
Total	660 46.09 100.00	772 53.91 100.00	1,432 100.00 100.00

*chi-square (4) = 30.332, p = 0.00

Split Decisions in Discharge Cases

Table 7 shows that discharge cases accounted for 1,432 decisions or approximately two-thirds of our aggregate data set. As noted above, most prior studies have reported case outcomes in terms of management wins, union wins, and split decisions. The size of our data set makes it possible to take a closer look at discharge case outcomes by focusing on the different types of remedies utilized in the split decision category. The two most common types of compromise remedies employed by arbitrators in discharge cases

are (1) reinstatement without back pay and (2) reduction to a suspension with partial back pay. Our study also identified a third, lesser-used compromise method, namely, an order entitling the grievant to a future vacancy.

Other Studies in Comparison. Three prior studies have examined specific remedial outcomes in discharge cases. The findings of these studies are depicted in Table 11.[22] They generally show that reinstatement without back pay constituted about two-thirds of all split discharge decisions with reduction to a suspension comprising the remaining one-third of those decisions. In another study analyzing 153 cases in which an arbitrator ordered reinstatement, Arthur Malinowski found that arbitrators ordered reinstatement without back pay three times more often than reinstatement with partial back pay.[23]

In contrast, Table 12 depicts our sample of specific remedial outcomes in discharge cases. Although our data generally fall within the range found in prior studies for full employer and full union wins, that is not the case with split outcomes. With respect to split decisions, two outlier findings are manifest. First, our overall share of split outcomes—27.79 percent—is well below the 29 to 38 percent range found in other studies. Second, reinstatement without back pay constitutes a smaller proportion of our split decisions than found in the previous studies, namely, 60.3 percent versus percentages that range from a low of 62.5 percent to a high, in Malinowksi's study, of about 75 percent. Indeed, our lower incidence of reinstatement without back pay outcomes ultimately accounts for both of these outlier findings.

---

[22]A reinstatement with partial or no back pay is generally consistent with our definition of a "split" decision. However, as previously noted, to these compromise outcomes we also add "entitlement to a future vacancy."

[23]Malinowski, *An Empirical Analysis of Discharge Cases and the Work History of Employees Reinstated by Labor Arbitrators*, 36 Arb. J. 31 (Mar. 1981).

TABLE 11: Studies of Remedy Outcomes

(in percent)

Discharge Study	Uphold the Discharge	Reinstate—Full Back Pay	Reinstate—Partial Back Pay	Reinstate—No Back Pay
Jones[1]	46.2	15.6	12.3	25.9
Jennings & Wolters[2]	46	22	12	20
Shearer[3]	51	20	8	21

[1]Jones, *Ramifications of Back-Pay Awards in Suspension and Discharge Cases*, in Arbitration and Social Change, Proceedings of the 22nd Annual Meeting, National Academy of Arbitrators, ed. Somers (BNA Books 1970), at 163 (analyzing 665 discharge cases reported in Labor Arbitration Reports for 1963–1968).

[2]Jennings & Wolters, *Discharge Cases Reconsidered*, 31 Arb. J. 164 (Sept. 1976) (reporting on 220 discharge cases appearing in Labor Arbitration Reports between 1971 and 1974).

[3]Shearer, *Reinstatement Without Back Pay—An Appropriate Remedy?*, 42 Arb. J. 47 (Dec. 1987) (examining 539 discharge cases reported in Labor Arbitration Reports for 1982–1986).

TABLE 12: Reinstatement Outcomes

(N-size and row % are the top and bottom value, respectively.)

Reinstatement Outcomes	Total
Uphold the Discharge	751 52.44
Reinstate—Full Back Pay	283 19.76
Reinstate—Partial Back Pay	149 10.41
Reinstate—Without Back Pay	240 16.76
Entitlement to Future Vacancy	9 0.63
Total	1,432 100.00

A Closer Look at Specific Types of Split Decisions: 1. Reinstatement Without Back Pay. Reinstatement without back pay is the most frequently occurring type of split decision, accounting for 16.76 percent of all discharge cases (240 occurrences). This outcome constitutes 35.7 percent of all cases in which reinstatement was ordered. Our data show that reinstatement without back pay was most often ordered (> 20 percent) in cases involving attendance or insubordination issues and least often ordered (< 13 percent) in cases involving dishonesty or performance issues.[24]

Reinstatement without back pay is the most controversial of the four principal remedial outcomes. A number of commentators have criticized the reinstatement without back pay option as imposing a remedy that does not replicate employer disciplinary practices.[25] A commonly voiced notion of arbitral jurisprudence is that an arbitrator's remedy should mirror that which an employer would have imposed if the employer had properly assessed the facts and circumstances of the case.[26] But, most employer disciplinary policies employ disciplinary suspensions up to only a maximum of a few weeks.[27] In contrast, a reinstatement without back pay order typically has the effect of an unpaid suspension of several months. Employers and unions alike look with suspicion on this compromise outcome. As summarized by John Teele in a 1964 article:

> The former is inclined to feel that if the man was guilty enough to be deprived of several months' back pay, then he should not have been returned to the payroll at all. The latter may be heard to argue that if the man was not guilty as charged, he should have been made whole.[28]

Our survey lends support to the underlying factual assertions made by these critics. Of the 516 cases in which employers imposed a disciplinary suspension, 47.5 percent (N = 245) of those cases resulted in a suspension of less than one week in duration and 52.5 percent (271) resulted in a suspension of more than one week. Although we did not code the precise length of employer-

---

[24]The offensive categories used in our study are described *supra* note 9.

[25]Zack, Arbitrating Discipline and Discharge Cases (Horsham, PA: LRP Publications 2000), at 148; Kienast, Reinstatement Without Back Pay and Back Pay Without Reinstatement: A Question of Appropriateness, 16 LERC Monograph Series 117, 118, 123 (2000); Shearer, *Reinstatement Without Back Pay—An Appropriate Remedy?*, 42 Arb. J. 48–49 (Dec. 1987); Teele, "*But No Back Pay is Awarded,*" 19 Arb. J. 103, 107–12 (June 1964).

[26]Kienast, *id.* at 118; Shearer, *id.* at 48–49.

[27]Kienast, *id.* at 118, 123.

[28]Teele, *supra* note 25 at 104.

imposed suspensions, these data suggest that the median length of employer-imposed disciplinary suspensions is about one week or five work days. In contrast, of the 240 cases in our sample in which an arbitrator reinstated a discharged employee without back pay, the mean period of lost pay equaled 8.67 months or about 37 weeks, which translates into a suspension of about 175 work days, assuming the suspended employee would work 20 days per month. As displayed in Figure 1, these outcomes range from a high mean of 14.08 months for off-the-job misconduct cases to a low mean of 6.62 months for drug/alcohol cases. Thus, our data suggest that the typical arbitrator-imposed order for reinstatement without back pay resulted in a period of denied pay 35 times longer than that which the typical employer would have imposed had it properly evaluated the circumstances resulting in discipline.

We also examined the impact of delay in reinstatement outcomes. The literature suggests two possible and somewhat contradictory effects in this regard. Some commentators have expressed the concern that arbitrators may use the reinstatement without back pay option more frequently and less defensibly in cases in which the length of time between the date of discharge and the date of the arbitrator's award is long.[29] The underlying notion, apparently, is that arbitrators are increasingly reluctant to saddle employers with the cost of back pay as that sum grows over time. Other commentators have suggested that arbitrators may be increasingly reluctant to order reinstatement in any form as the length of time between the date of discharge and date of award expands. In this regard, a study undertaken by Nelson and Uddin found that the likelihood that a grievant will be reinstated in a discharge case generally declined by 8.6 percent for each additional month of delay leading up to the arbitral decision.[30]

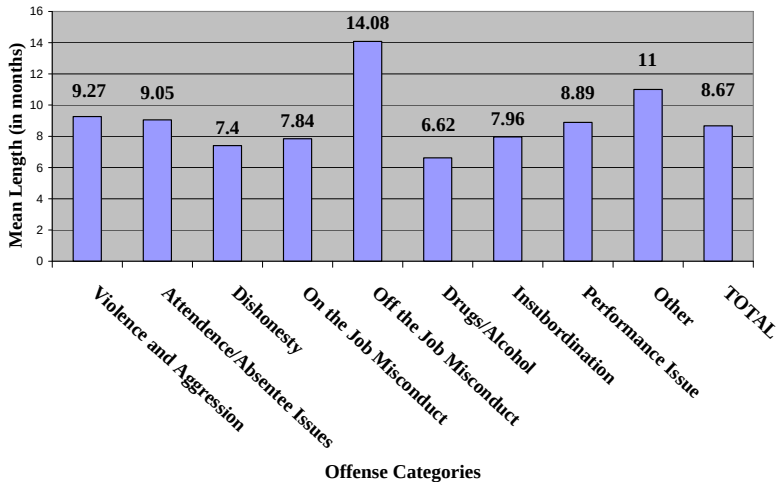
Table 13 shows selected outcomes by seven time-specific categories, each reflecting the length of time between a grievant's date of discharge and the date of the corresponding arbitration award (hereafter referred to as "delay"). Specifically, for each delay category, we identify (1) the number of all discharge cases that are upheld, (2) the number that resulted in the grievant's reinstate-

---

[29] *See, e.g.*, Saxton, Miller & Fallon, *The Discipline and Discharge Case: Two Devil's Advocates on What Arbitrators are Doing Wrong*, in Arbitration of Subcontracting and Wage Incentive Disputes, Proceedings of the 32nd Annual Meeting, National Academy of Arbitrators, eds. Stern & Denis (BNA Books 1980), at 63, 84–85.

[30] Nelson & Uddin, *The Impact of Delay on Arbitrators' Decisions in Discharge Cases*, 3 Lab. Stud. J. 3, 12 (1998).

Figure 1. Mean Length of No Back Pay Period by Offense Categories



ment without back pay, and (3) the number that resulted in the grievant's reinstatement under any condition (i.e., with full, partial, or no back pay). Using percentage or share data, the separate graphs of the relationship between reinstatement outcomes and delay tell a slightly more nuanced story. See Figure 2. First, the bottom graph in Figure 2 is "flat," which is dismissive of the first hypothesis discussed above: namely, that delay and reinstatement without back pay are positively related. Second, the top graph in Figure 2 suggests that cases that make it to arbitration after 18 months correlate positively with the discharge decision being upheld.[31] This disaggregated glimpse at the data lends modest but inconclusive support to the second hypothesis that delay and reinstatement are inversely related. Ultimately, we plan to examine

[31]The issuance of a reinstatement award was 29% less likely in those cases in which the award followed the discharge by more than two years as compared with those awards issued following a span of only 10 to 12 months. On the other hand, the number of awards ordering reinstatement without back pay fell by an even larger proportion (37%) over that same span. These findings suggest that arbitrators in delayed proceedings are more greatly influenced by a reluctance to reinstate long-departed employees than they are by a motivation to spare employers the rising cost of a back pay award.

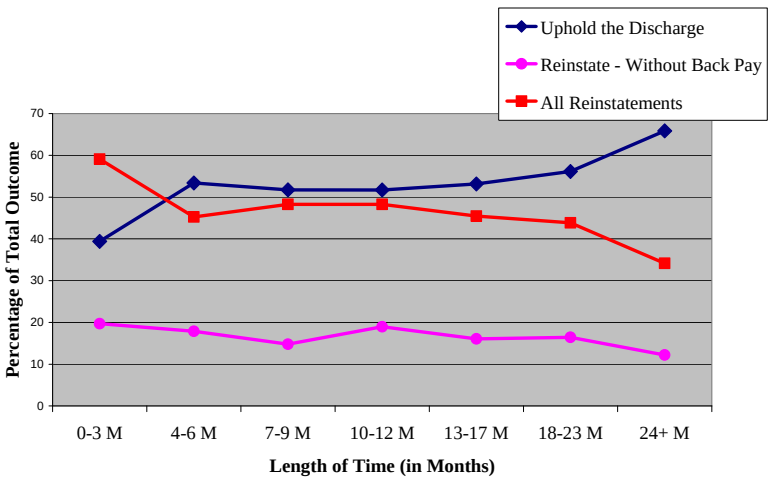
TABLE 13: Delay and Reinstatement Outcomes*

(N-size and row % are the top and bottom value, respectively.)

Delay (in months)	Discharge Upheld	Reinstate—No Back Pay	All Reinstatements	Total Decisions
0–3	26 40.00	13 20.00	39 60.00	65
4–6	236 54.13	79 18.12	200 45.87	436
7–9	224 51.73	64 14.78	209 48.27	433
10–12	120 51.50	44 18.88	113 48.50	233
13–17	76 53.90	23 16.31	65 46.10	141
18–23	41 56.16	12 16.44	32 43.84	73
24+	27 65.85	5 12.20	14 34.15	41
Total Columns	750 52.74	240 16.88	672 47.26	1,422

*The column categories of the outcome variable are not mutually exclusive.

FIGURE 2. Outcomes per Variation in Length of Time from Discharge to Date of Award

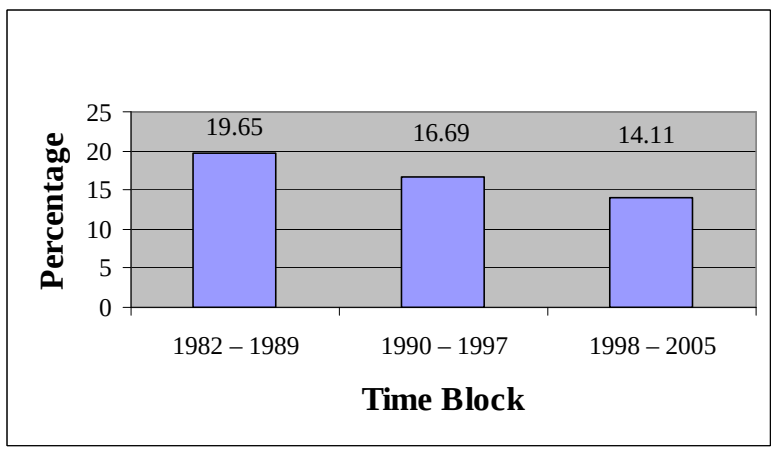


the partial relationship between outcomes and delay within a multivariate analytic framework.

Another matter warranting closer inspection is the reduced prevalence of reinstatement without back pay outcomes in our data. As noted in Table 11, earlier studies generally showed that reinstatement without back pay outcomes comprised between 20 and 26 percent of all discharge cases. Our findings, in contrast, exhibit a much lower 16.76 percent relative frequency.

One possible explanation for this discrepancy is that published awards contain more reinstatement without back pay outcomes than awards in general because publishers think that those borderline cases will be more interesting to their readership. Another possible explanation is that the incidence of reinstatement without back pay outcomes may be declining in response to the persistent drumbeat of criticism. If that were the case, one would expect to see a decline in the proportion of reinstatement without back pay outcomes over time in our data base. In fact, our findings bear out such a trend. Figure 3 shows the percentage of reinstatement without back pay outcomes in the three 8-year increments spanning the years between 1982 and 2005, inclusively. From a 19.65 percent prevalence in the 1982–1989 set of decisions—not far below the range found in other contemporary studies—the Minnesota

Figure 3. Prevalence of Reinstatement without Back Pay



proportion of reinstatement without back pay outcomes fell to 16.69 percent in 1990–1997 and still further to 14.11 percent in 1998–2005.[32] Indeed, a comparison of the first and third periods reveals a significant 28.2 percent drop in reinstatement without back pay outcomes. Whether this finding foretells an ongoing shift in remedy outcomes remains to be seen.

A Closer Look at Specific Types of Split Decisions: 2. Reduction to Suspension. Our data show that 10.41 percent of all discharge cases, consisting of 149 decisions, involved instances in which an arbitrator reinstated the grievant and reduced the discharge to a period of suspension without pay. This amounts to 22.2 percent of all reinstatement outcomes, representing the least prevalent of the three reinstatement options.

The two most likely offenses to result in a reduction to suspension are on-the-job misconduct and insubordination. With respect to these offenses, a plausible explanation for this result is that arbitrators are most prone to concur in management's allegation that the grievant engaged in the alleged misconduct, while nonetheless concluding that management overreacted in selecting discharge as the appropriate level of discipline.

A most remarkable finding regarding the reduction to suspension option concerns the length of the suspension period. The mean length of the arbitrator-ordered suspension period in our sample is 61.04 work days. Even after deleting an 849-day suspension outlier from the analysis, the mean length of the remaining suspensions is 55.57 days, or approximately 11 work weeks.[33] The mean suspension length by offense (depicted in Figure 4) ranges from a high of approximately 75 work days for on-the-job misconduct to a low of approximately 25 work days for dishonesty. The mode outcome, consisting of more than one-sixth of all reduction to suspension results, is a 30-day suspension period.

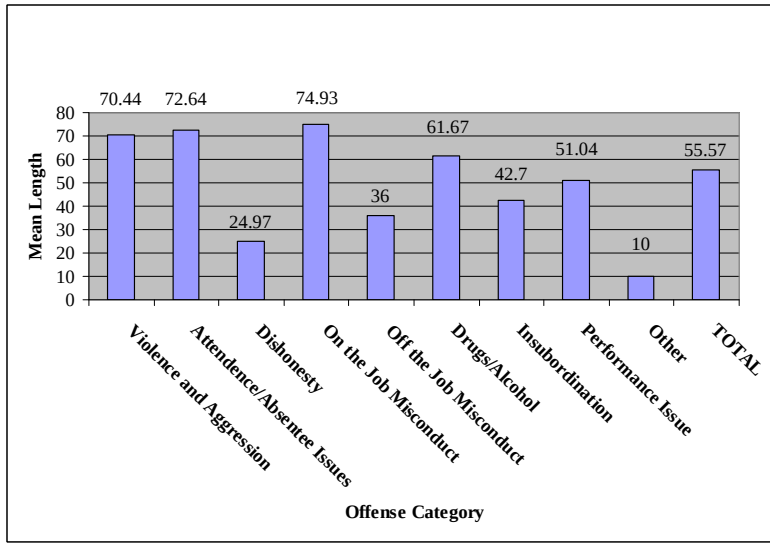
These findings demonstrate that the reduction to suspension outcomes also do not replicate employer-imposed suspension lengths. Although the mean reduction to suspension period is less than one-third the length of the mean reinstatement without back pay period, it still is 11 times the length of the typical one week employer-imposed suspension. This suggests that many

---

[32]The test statistic for the association pictured in Figure 3 is chi-square (2) = 4.4708, p = .107.

[33]We also removed four other reductions to suspension from this pool because of the indeterminate nature of the suspension period. As a result, our figures are based on a data set consisting of the remaining 144 reduction to suspension outcomes.

Figure 4. Mean Length (in days) of Suspension Imposed by Offense Category and Total



arbitrators are determining the length of the reduction to suspension period on the basis of personal notions of equity as opposed to a correlation with employer practices. As such, it appears warranted to conclude that many arbitrators use the reduction to suspension option to craft a second type of compromise outcome similar to that for which the reinstatement without back pay option is so often criticized.

A Closer Look at Specific Types of Split Decisions: 3. Entitlement to a Future Vacancy. Our survey encountered an unexpected third category of split decisions. Nine decisions, or 0.63 percent of all discharge cases, resulted in an order awarding the grievant not reinstatement, but rather the opportunity to fill a future vacancy. Although this was a rare outcome, the surprising fact is that it happened as often as it did. This is not an outcome that is referenced in the arbitration literature or that has been considered by prior studies. This remedy was most often invoked by arbitrators in cases involving violence or poor performance. Six of the nine instances were ordered by less experienced, non-NAA arbitrators.

Seven Tests and Quantum of Proof

The Seven Tests

Arbitrators adjudicating employee discipline cases are asked to determine, in the language of the parties' collective bargaining agreements, whether the employee has been discharged or disciplined for "just cause."[34] Because contracts typically do not further specify the meaning of "just cause," advocates for employers or unions necessarily must present their evidence and arguments, and arbitrators their analysis, within a more defined understanding of the concept. Although countless presentations at NAA meetings and other publications have addressed various aspects of the phrase, no single articulation of the requirements of just cause has gained the salience of "The Seven Tests of Just Cause," first articulated by Arbitrator Carroll R. Daugherty in the 1960s.[35] One commentator describes it as "the single most definitive statement of just cause."[36] The Academy's *Common Law of the Workplace* describes the Seven Tests as "undeniably influential."[37] A 560-page treatise on just cause, now in its third edition, is entirely structured around application of the Seven Tests.[38] The Seven Tests are widely used in materials designed for the training of arbitrators and labor arbitration advocates.[39]

The Seven Tests are posed in the form of questions. In this articulation, a "no" answer to any one of the questions "normally signifies that just and proper cause did not exist."[40] The questions are:

---

[34]One survey found the requirement of "just cause" or "cause" for employee discipline included in 92 percent of collective bargaining agreements. Basic Patterns in Union Contracts, 14th ed. (BNA Books 1995), at 7. Arbitrators and courts have found the requirement implicit in contracts without an explicit "just cause" provision. *See, e.g., SFIC Properties, Inc. v. Machinists, District Lodge 94*, 103 F.3d 923 (9th Cir. 1996) (just cause implicit "in all modern day collective bargaining agreements").

[35]The Seven Tests were first listed in *Grief Bros. Cooperage Corp.*, 42 LA 555 (Daugherty 1964), and further explained in *Enterprise Wire Co.*, 46 LA 359 (Daugherty 1966).

[36]Ver Ploeg, *Investigatory Due Process and Arbitration*, in Arbitration 1992: Improving Arbitral and Advocacy Skills, Proceedings of the 45th Annual Meeting, National Academy of Arbitrators, ed. Gruenberg (BNA Books 1993), at 220, 223–24.

[37]St. Antoine, ed., The Common Law of the Workplace: The Views of Arbitrators, 2d ed. (BNA Books 2005), at §6.12 Comment.

[38]Koven & Smith, Just Cause: The Seven Tests, 3d ed. (BNA Books 2006).

[39]Dunsford, *Arbitral Discretion: The Tests of Just Cause*, in Arbitration 1989: The Arbitrator's Discretion During and After the Hearing, Proceedings of the 42nd Annual Meeting, National Academy of Arbitrators, ed. Gruenberg (BNA Books 1990), at 23, 33 (used in training materials of the American Arbitration Association and American Bar Association).

[40]*Grief Bros. Cooperage Corp.*, 42 LA 555, 557 (Daugherty 1964).

1. Did the company give to the employee forewarning or foreknowledge of the possible or probable disciplinary consequences of the employee's conduct?
2. Was the company's rule or managerial order reasonably related to (a) the orderly, efficient, and safe operation of the company's business and (b) the performance that the company might properly expect of the employee?
3. Did the company, before administering discipline to an employee, make an effort to discover whether the employee did in fact violate or disobey a rule or order of management?
4. Was the company's investigation conducted fairly and objectively?
5. At the investigation, did the "judge" obtain substantial evidence or proof that the employee was guilty as charged?
6. Has the company applied its rules, orders, and penalties evenhandedly and without discrimination to all employees?
7. Was the degree of discipline administered by the company in a particular case reasonably related to (a) the seriousness of the employee's proven offense and (b) the record of the employee in his service with the company?

The Seven Tests were initially offered as a description of what arbitrators actually did, but they have been criticized both as a matter of descriptive accuracy and of theory.[41] Although several commentators have expressed some doubt about whether arbitrators actually assess just cause by using the Seven Tests, we are aware of no previous empirical study that sought to determine the proportion of arbitration awards that rely upon the Seven Tests.[42]

As we test the influence of the Seven Tests, we should be clear about what we are measuring and what we are not. We are not

---

[41]Clarke, *To What Extent Do and Should the Seven Tests Guide Arbitrators or the Parties?*, in Arbitration 2002: Workplace Arbitration: A Process in Evolution, Proceedings of the 55th Annual Meeting, National Academy of Arbitrators, ed. Coleman (BNA Books 2003), at 51; Dunsford, *supra* note 39; Ver Ploeg, *Investigatory Due Process.* An alternative, and more thoughtful, formulation of the meaning of just cause, was articulated in Abrams & Nolan, *Toward a Theory of 'Just Cause' in Employee Discipline Cases*, 85 Duke L.J. 594 (1985).

[42]The closest to an empirical examination was Professor Ver Ploeg's reading of 10 years of published arbitral decisions from which she concluded that there was no consensus among arbitrators on the question of whether an employer's investigatory procedural irregularities should invalidate the employer's discipline. Ver Ploeg, *Investigatory Due Process,* 228. Professor Ver Ploeg, however, provided no tally of the cases taking alternative perspectives.

assessing whether some of the inquiries posed by the Seven Tests are, individually, relevant to the arguments of the parties and the decisions of arbitrators, as indeed they surely are. For example, as noted in Table 1, arbitrators do mitigate penalties when an employee's discipline appears disproportionate to the offense and when the discipline appears excessive in light of the employee's long years of service. Rather, we are examining, to the extent possible, whether the rubric of the Seven Tests, as a unified defining formulation of the meaning of just cause, explicitly manifests itself in the arguments of parties and the awards of arbitrators.

In our study, we coded the answer to the question: "Does the arbitrator draw on Daugherty's Seven Tests in reaching the arbitrator's 'guilt' and/or 'just cause' determinations?" As seen in Table 14, in 91.44 percent of decisions the arbitrator did *not* rely on the Seven Tests; in only 8.56 percent of decisions did arbitrators explicitly use that structure for analysis of just cause.

We also sought to assess the extent to which the parties used the Seven Tests and the extent to which such advocacy affected the arbitrators' reliance on that standard. As arbitrators' awards typically summarize the parties' arguments, we were able to measure whether there was any mention in the arbitrator's award of either the employer or the union invoking the Seven Tests.[43] As shown in Table 14**,** using that measure**,** we see that neither the union nor the employer advocate referenced the Seven Tests in 96.11 percent of the cases. Arbitrators were much more likely to employ the Seven Tests in cases in which advocates had relied upon it. That suggests that much of the time arbitrators rely on the Seven Tests they are doing so to accommodate the perspective of the advocates rather than because the arbitrators independently believed that the Seven Tests was the appropriate method of analysis. On the other hand, arbitrators readily rejected the parties' reliance on the Seven Tests in writing their awards. Looking at only those cases in which the arbitrator's award noted that one or both advocates relied on the Seven Tests, arbitrators were nearly twice as

[43]As we did not have available the parties' briefs or oral arguments, our measure of the extent to which the parties invoked the Seven Tests is likely understated.

Table 14: Use of the Seven Tests by Advocates and Arbitrators*

(N-size, row %, and column % are the top, middle, and bottom value, respectively.)

	Arbitrator Used 7 Tests	Arbitrator Did Not Use 7 Tests	Row Totals
One or Both Parties Invoked 7 Tests	28 35.00 15.91	52 65.00 2.77	80 100.00 3.89
Neither Party Invoked 7 Tests	148 7.49 84.09	1,827 92.51 97.23	1,975 100.00 96.11
Column Totals	176 8.56 100.00	1,879 91.44 100.00	2,055 100.00 100.00

*chi-square (1) = 74.2839, p = 0.000

likely *not* to use the Seven Tests in reaching their conclusions (65 percent of the time) as to adopt the advocates' approach (35 percent of the time).

Our survey of 2,055 decisions included only 36 cases (1.75 percent) in which an award mentioned an employer's reliance on the Seven Tests and only 56 cases (2.73 percent) in which union reliance was noted. Unions might be more likely than employers to invoke the Seven Tests because its analysis directs an arbitrator to overturn employee discipline, despite the employee's guilt of the alleged misconduct, because of the employer's procedural failing. As shown in Table 15, whether an attorney or a non-attorney advocates on behalf of a party does not significantly affect the likelihood of the advocate relying upon the Seven Tests.

TABLE 15: Association: Attorney Representation and Invocation of the Seven Tests*

(N-size, row %, and column % are the top, middle, and bottom value, respectively.)

Attorney Representation	Advocate Invoked 7 Tests	Advocate Did Not Invoke 7 Tests	Row Totals
Union—Attorney	30 2.87 53.57	1,017 97.13 50.88	1,047 100.00 50.95
Union–No Attorney	26 2.58 46.43	982 97.42 49.12	1,008 100.00 49.05
Column Totals* *chi-square (1) = .1584, p = .691	56 2.73 100.00	1,999 97.27 100.00	2,055 100.00 100.00
Employer—Attorney	24 1.71 66.67	1,383 98.29 68.50	1,407 100.00 68.47
Employer—No Attorney	12 1.85 33.33	636 98.15 31.50	648 100.00 31.53
Column Totals* *chi-square (1) = .055, p = .815	36 1.75 100.00	2,019 98.25 100.00	2,055 100.00 100.00

Moreover, although union advocates, both attorneys and non-attorneys, were more likely than employer advocates to invoke the Seven Tests, as shown in Table 16, when arbitrators did rely on the Seven Tests there was no significant improvement in the union's likelihood of prevailing, particularly its likelihood of an arbitrator finding no just cause for any discipline.

Cases decided by non-members of the NAA were nearly twice as likely as cases decided by Academy members to rely upon the Seven Tests, as indicated in Table 17. Indeed, the association

Table 16: Association: Use of the Seven Tests and Outcomes*

(N-size and row % are the top and bottom value, respectively.)

Arbitrator Used Seven Tests	Found Just Cause for Discipline	Found No Just Cause for Discipline	Found Just Cause for Lesser Discipline	Total
Yes	93 52.84	39 22.16	44 25.00	176 8.56
No	929 49.44	402 21.39	548 29.16	1,879 91.44
Total	1,022 49.73	441 21.46	592 28.81	2,055 100.00

*chi-square (2) = 1.3867, p = 0.500

Table 17: Association: NAA Membership and Use of the Seven Tests*

(N-size and row % are the top and bottom value, respectively.)

Arbitrator Used Seven Tests	Arbitrator is NAA Member	Arbitrator is Not NAA Member	Total
Yes	63 5.88	113 11.50	176 8.56
No	1,009 94.12	870 88.50	1,879 91.44
Total	1,072 100.00	983 100.00	2,055 100.00

*chi-square (1) = 20.6714, p = 0.000

between NAA membership and use of the Seven Tests is statistically quite strong.

Quantum of Proof. It appears to be universally accepted in labor arbitration that in employee discipline cases the employer bears the burden of proving that the employee was disciplined for just

cause, rather than the union having to prove that the employer's discipline was not for just cause.[44] It has, however, long been noted that arbitrators differ on the extent of the employer's burden, that is, whether to apply the "preponderance of the evidence" standard used for most issues in litigated civil cases, whether to apply the "beyond a reasonable doubt" standard used in criminal law, or whether to apply the intermediate standard of "clear and convincing evidence" used for reasons of policy for a limited number of specific litigated issues.

Some arbitration commentators and arbitrators reject the entire question of quantum of proof as "just playing games with words"[45] or as a "distracting legalism."[46]

In the *Common Law of the Workplace*, the Academy sought to articulate "some generally accepted approaches toward commonly encountered problems."[47] With regard to the quantum of proof applied by arbitrators in discipline cases, §6.10 of the *Common Law of the Workplace* includes the following assertions:

(a) "For *most* arbitrators, the normal quantum of proof required in disciplinary cases is 'preponderance of the evidence.'"
(b) "For a *minority*, it is 'clear and convincing evidence.'"
(c) In cases involving allegations of an offense that would be "a serious breach of law or would be viewed as moral turpitude" "*most* arbitrators require a higher quantum of proof, typically expressed as 'clear and convincing evidence.'"
(d) In such cases of alleged serious misconduct, "*some*" arbitrators would employ a "beyond a reasonable doubt" standard, but *most* conclude that such a high standard has "no place" in labor arbitration.

Prior to our study, however, there had been scant empirical testing of these assertions. The few studies previously conducted focused on whether heightened standards are used in cases alleging crimes or other matters of moral turpitude. In a survey, reported by Jeffrey Small and J. Timothy Spreche in 1984, more than 1,000 arbitrators were asked what standard of proof they would apply

---

[44]Common Law of the Workplace, *supra* note 2, at §6.9.
[45]Remarks of Feller, *Admissibility of Evidence,* in Arbitration 1982: Conduct of the Hearing, Proceedings of the 35th Annual Meeting, National Academy of Arbitrators, eds. Stern & Dennis (BNA Books 1983), at 136.
[46]Common Law of the Workplace, *supra* note 2 at §6.10 Comment.
[47]*Id.*, Preface to First Edition, ix.

in a hypothetical case alleging theft. Of the arbitrators surveyed, 34.3 percent said that they would use "clear and convincing evidence" and 22.7 percent said they would use "beyond a reasonable doubt."[48] A 1989 study by Kenneth W. Thornicroft explored the quantum of proof used in 145 published cases from 1985–1988, involving employees discharged for drug or alcohol use. He found that 35 percent of the drug cases and 47 percent of the alcohol cases employed the "clear and convincing" standard. Thornicroft said that 9 percent of the drug cases and 14 percent of the alcohol cases used "beyond a reasonable doubt."[49]

We were able to identify or infer the standard of proof applied by the arbitrator in all 2,055 cases in our data base. In 1,609 cases, or 78.30 percent, arbitrators did not state the quantum of proof that they were applying. In 200 cases, or 9.73 percent, arbitrators explicitly stated that they were applying the standard of "preponderance of the evidence." We assume that, when arbitrators do not overtly state what quantum of proof they are applying, they are determining whether the employer's claim of just cause is more likely true than the union's claim of lack of just cause. Thus, we are assuming that an unstated quantum of proof is equivalent to a stated standard of "preponderance of the evidence." Given this assumption, arbitrators applied the "preponderance" standard in 88.03 percent of decisions analyzed. (See Table 18.) Our findings, then, support the assertion of the *Common Law of the Workplace* that in most cases arbitrators assess just cause by a standard of "preponderance of the evidence" and that only a minority employ a standard of "clear and convincing evidence."

We can also test the assertion in the *Common Law of the Workplace* that in cases alleging conduct that could constitute crimes or other matters of moral turpitude, "most arbitrators require a higher quantum of proof, typically expressed as 'clear and convincing evidence.'" With respect to employees accused of offenses that might be viewed as crimes or moral turpitude,[50] our analysis and the chi-square reported in Table 19 contradict the expectation of the *Common Law of the Workplace* that most arbitrators require a higher quantum of proof in such cases. Indeed, arbitra-

---

[48]Small & Spreche, Report of American Arbitration Association Survey of Labor Arbitrators, 1984 Daily Lab. Rep. (BNA) (Dec. 5), No. 234: E-14.

[49]Thornicroft, *Arbitrators and Substance Abuse Discharge Cases: An Empirical Assessment*, Lab. Stud. J. 40, Table at 58 (Winter 1989).

[50]Offenses classified as "off-the-job misconduct" are included here as an employer would be likely to consider disciplining an employee for misconduct away from the workplace only when it was of a particularly serious nature.

TABLE 18: Quantum of Proof

(N-size and row % are the top and bottom value, respectively.)

Standard	N	%
Preponderance of the Evidence or No Specific Standard Stated	1,809	88.03
Clear and Convincing Evidence	207	10.07
Beyond a Reasonable Doubt	39	1.90
Total	2,055	100.00

TABLE 19: Quantum of Proof Correlated with Offense Category*

(N-size and row % are the top and bottom value, respectively.)

Alleged Offense Category**	Preponderance of the Evidence or No Specific Standard Stated	Clear and Convincing Evidence	Beyond a Reasonable Doubt	Total**
Offenses Involving Crimes or Moral Turpitude***	868 83.95	133 12.86	33 3.19	1034 100.00
Offenses Not Involving Crimes or Moral Turpitude***	1269 91.29	112 8.06	9 0.65	1390 100.00
Total**	2,137	245	42	2,424

*chi-square (2) = 39.3245, $p < .0001$

**Some cases involved more than one offense category (e.g., the discipline might have been based on both a performance issue and on insubordination).

***Offenses categorized as involving crimes or moral turpitude include dishonesty, off-the-job misconduct, violence and aggression, and drugs and alcohol. Offenses classified as not involving crimes or moral turpitude include on-the-job misconduct, insubordination, performance issues, attendance/absenteeism issues, and offenses classified as "other."

tors only slightly less often applied the standard of "preponderance of the evidence" in cases alleging serious misconduct (83.95 percent) than in our data as a whole (88.03 percent). Our raw data do, however, support the conclusion of the *Common Law of the Workplace*, as well as the Thornicroft study and Small-Spreche survey, that in cases alleging serious offenses, when arbitrators do apply a heightened standard of proof, they are much more likely to choose "clear and convincing" (12.86 percent) rather than "beyond a reasonable doubt" (3.19 percent).

Table 20 compares the standard of proof by arbitrators in each category of employee offense. Here we see that arbitrators were most likely to apply a heightened standard of proof in cases involving allegations of dishonesty and least likely to apply a heightened standard in cases involving attendance, insubordination, and performance issues. Table 20's variables of offense and standard of proof are statistically related.

The clearest indication that "most arbitrators" are not applying a heightened standard of proof in cases in which employees are alleged to have committed the most serious kinds of misconduct comes from reviewing those cases in which arbitrators noted in their awards that the employee had been charged with a crime in the criminal justice system for the conduct for which the employee was disciplined by the employer. Table 21 indicates that when arbitrators considered whether there was just cause for an employer to discipline an employee for the same conduct that had given rise to a criminal charge, arbitrators applied a heightened standard of proof less than one-quarter of the time (23.96 percent). Not surprisingly though, the data do indicate that arbitrators are more likely to apply a heightened standard of proof in cases of alleged criminality than in cases in which no crime was alleged. (The chi-square test for Table 21 shows that the two variables of alleged criminality and standard of proof are dependent.) Even in cases of alleged criminality, however, arbitrators were nearly twice as likely to use the "clear and convincing" standard (15.63 percent) than that of proof "beyond a reasonable doubt" (8.33 percent).

TABLE 20: Quantum of Proof Correlated with Offense Category*

(Arranged in descending order of likelihood of heightened quantum of proof; N-size and row % are the top and bottom value, respectively)

Alleged Offense Category**	Preponderance of the Evidence or No Specific Standard Stated	Clear and Convincing Evidence	Beyond a Reasonable Doubt	Total**
Dishonesty	309 81.96	52 13.79	16 4.24	377 100.00
Off-the-Job Misconduct	79 84.04	10 10.64	5 5.32	94 100.00
Violence & Aggression	397 85.01	62 13.28	8 1.71	467 100.00
Drugs/Alcohol	83 86.46	9 9.38	4 4.17	96 100.00
On the Job Misconduct	157 89.71	16 9.14	2 1.14	175 100.00
Insubordination	428 91.26	38 8.10	3 0.64	469 100.00
Performance Issue	338 91.35	30 8.11	2 0.54	370 100.00
Attendance/ Absentee Issues	323 91.76	27 7.67	2 0.57	352 100.00
Other	23 95.83	1 4.17	0 0.00	24 100.00
Total**	2,137	245	42	2,424

* chi-square (16) = 54.0001, p < .0001
**Some cases involved more than one offense category (e.g., the discipline might have been based on both a performance issue and on insubordination).

In Table 22, we analyze the relationship between the quantum of proof applied by the arbitrator and case outcomes. As expected, employers were most likely to prevail in full in those cases in which arbitrators imposed on the employer the responsibility of persuasion by the lowest standard, "preponderance of the evidence" (51.08 percent). Further, arbitrators were most likely to

Table 21: Quantum of Proof in Cases in Which Employees Were Charged With a Crime in the Criminal Justice System*

(N-size and row % are the top and bottom value, respectively.)

Quantum of Proof	Employee Charged With Crime for Conduct Alleged	Employee Not Charged With Crime for Conduct Alleged
Preponderance of the Evidence or No Specific Standard Stated	73 76.04	1,736 88.62
Clear and Convincing Evidence	15 15.63	192 9.80
Beyond a Reasonable Doubt	8 8.33	31 1.58
Total	96 100.00	1,959 100.00

*chi-square (4) = 37.2368, p < 0.000

Table 22: Association: Quantum of Proof and Outcomes*

(N-size and row % are the top and bottom value, respectively}

Quantum of Proof	Just Cause for Employer's Discipline	Just Cause for Lesser Discipline	No Just Cause for Any Discipline	Total Rows
Preponderance of the Evidence or No Specific Standard Stated	924 51.08	520 28.75	365 20.18	1,809 100.00
Clear and Convincing Evidence	82 39.61	65 31.40	60 28.99	207 100.00
Beyond a Reasonable Doubt	16 41.03	7 17.95	16 41.03	39 100.00
Total Columns	1,022 49.73	592 28.81	441 21.46	2,055 100.00

*chi-square(4) = 21.4046, p = 0.000

find that employers had no just cause for imposition of any discipline in those cases in which arbitrators required the highest level of evidence, that of proof "beyond a reasonable doubt" (41.03 percent). The "beyond a reasonable doubt" standard was least likely to be associated with the conclusion that the employee's conduct warranted just cause for lesser discipline than that imposed by the employer (17.95 percent). That makes sense in light of the category of cases in which the "beyond a reasonable doubt" standard is most likely to be used—those with allegations of the most serious kind of misconduct. In such cases, if the employee is found to have committed the conduct, just cause is likely to be found, but if the employee is found not to have been the perpetrator, then no discipline would be warranted. Arbitrators would be least likely to find lesser discipline appropriate if an employee actually is found to have committed a very serious offense. The chi-square test indicates that is it highly probable that the quantum of proof and outcomes are dependent.

Conclusions

This paper provides a preliminary assessment of what may be the most complete collection of a state's published and unpublished discipline and discharge arbitration decisions ever subject to systematic analysis. The size of the data base and the large variety of coded survey items about both the cases and the arbitrators who decided them permits empirical testing of many assertions in the arbitration literature about the nature of discipline and discharge decisionmaking: Assertions that have been drawn mainly from informal reviews of published decisions or, at best, from empirical studies, nearly all of which were based on unrepresentative published decisions.

In some respects, our findings support the conclusions in the literature about arbitral decisionmaking and in the empirical studies on which some of those conclusions were based, but in other respects our results challenge those statements and studies. Our findings, for example, are consistent with prior generalizations about arbitrator fidelity to progressive discipline and that arbitrators tend to enforce last-chance agreements. In addition, we find that the most controversial type of split remedial decision, reinstatement without back pay, was the most frequently occurring kind of split decision. We confirm the findings of other studies that split decisions result in suspensions without pay far longer

than would be imposed by employers as a disciplinary sanction. We also cautiously affirm the role of delay in reducing a grievant's chance of reinstatement.

On the other hand, our data give reason to question the assumption that arbitrators typically retain jurisdiction when they issue back pay awards. We observe the Seven Tests being far less influential in actual assessment of just cause than the literature would suggest. Our results also indicate that arbitrators invoke heightened proof requirements, such as "clear and convincing evidence" and "beyond a reasonable doubt" much less frequently than claimed in the literature or found by prior studies, even in cases where employees have been charged with criminal conduct.

This presentation is the beginning of our efforts to draw from our newly compiled data set the most empirically valid picture of the nature of decisionmaking in discipline and discharge labor arbitration yet attempted. We plan to continue these explorations in a future book and subsequent articles that will also include the application of more sophisticated statistical techniques.

II. How and Why Labor Arbitrators Decide Discipline and Discharge Cases: An Empirical Exmination—Comments

Theodore J. St. Antoine*

Laura Cooper, Mike Bognanno, and Steve Befort (CBB) have made a major contribution to our understanding of the decisional process in the arbitration of discipline and discharge cases. Nels Nelson has carefully examined their methodology and the reasons their study holds so much greater potential for drawing sound conclusions than previous efforts based on considerably more limited data bases. My comments will deal with a particular theme that runs through the study: the extent to which it confirms or challenges a number of the generalizations about arbitrators' views set forth in the NAA's *The Common Law of the Workplace*, published in 1998 and revised in 2005.[1]

*Member, National Academy of Arbitrators, Degan Professor Emeritus of Law, University of Michigan.

[1]St. Antoine, ed., The Common Law of the Workplace: The Views of Arbitrators, 2d ed. (BNA Books 2005).

Let me first say a word about the origins, philosophy, and modus operandi of *The Common Law*, because I think they say a lot about the credibility of the resulting product. The work was the brainchild of Arnold Zack, Academy President in 1994–95, with his principal co-conspirators being his successor Presidents, Ted Weatherill and George Nicolau. *The Common Law* was designed to commemorate the 50th Anniversary of the NAA in 1997 by summing up some of the leading arbitral principles developed over its first half century. Many veteran arbitrators rarely published their decisions and media attention often focused on the sensational case rather than the more typical. Although there had been good encyclopedic treatments, it was felt that a fairly short, authoritative overview would be more useful to the less experienced arbitrator or advocate.

Arnie caught me in a weak moment, enjoying a light teaching load while visiting at Cambridge University, and I agreed to be editor of the volume. We next rounded up 15 star performers among Academy members to do the hard job of writing on everything from arbitral practice and procedure to remedies in arbitration. And then, in a move I considered very important, we enlisted the aid of an advisory group of eight former Academy Presidents, chaired by Dick Mittenthal. They went over at least one chapter each, and did not hesitate to criticize and suggest improvements. The 15 writers and I engaged in numerous debates over drafts covering the more sensitive areas. Selected portions of early drafts dealing with some of the most controversial issues were placed before the entire membership attending three different general meetings as well as some of the regional meetings of the Academy. When there was respectable support for different positions on certain issues, we would include the various points of view. The upshot, I believe, reflected arbitral thinking in the United States and Canada as accurately as could be ascertained in the judgment of a highly able and experienced group of arbitrators. But, of course, it did not constitute the sort of actual head count we have before us today.

What, then, do the comparisons have to tell us? As might be expected, the empirical study supported the validity of a number of the generalizations contained in *The Common Law*. Thus, as seen in decisions reducing employer discipline, "most" arbitrators do adopt the principle of progressive discipline, even though a particular collective bargaining agreement may say nothing about the subject. Not surprisingly, arbitrators "generally" enforce last-

chance agreements. On the famous (or infamous) Seven Tests of Arbitrator Carroll Daugherty for "just cause," CBB report that in only about 9 percent of the decisions they studied did arbitrators rely on them. *The Common Law* concedes that the Seven Tests have been "influential," but cites a "critically convincing" alternative view by Professor John Dunsford. Naturally, the areas I find most interesting are those in which the empirical study seems to refute the positions taken by *The Common Law*.

For example, there may be a conflict concerning the retention of jurisdiction when the arbitrator determines that back pay is appropriate and remands the case to the parties to compute it. *The Common Law* states: "Such a remedy is *usually*, but not always, accompanied by a retention of jurisdiction by the arbitrator in the event that there is a subsequent dispute over the amount" (emphasis added). The CBB empirical study indicated that in 432 cases where the employee was reinstated with full or partial back pay, the arbitrator did not retain jurisdiction about 73 percent of the time. But the precise question asked of the data was whether the arbitrator "expressly" retained jurisdiction to resolve any subsequent disputes. That is a big difference. Although I invariably retain jurisdiction in such situations in an ad hoc appointment, I do not feel any such need when I am handling a case as a permanent arbitrator or as a member of a board of arbitrators. There the understanding of everyone is that the issue will come back for a final ruling if the parties are not able to resolve it. In addition, an entirely plausible argument can be made that an award that provides for back pay but does not compute it and does not deal with such natural additional questions as mitigation of damages is actually an interim award, and must be regarded as implicitly retaining jurisdiction if the award is not to be treated as defective for lack of completeness and finality. At any rate, I think it deserves further inquiry whether, at least in such a common situation as the resolution of computation issues in back pay cases, the arbitrator and the parties assume that retention of remedial jurisdiction is implicit. Incidentally, at the Academy's Business Meeting today a motion was adopted to modify the Code of Professional Responsibility to permit an arbitrator to retain remedial jurisdiction even over the objection of one of the parties.

The question of the quantum of proof required in discipline and discharge cases is another area in which the generalizations of *The Common Law* may be at odds with the findings of the CBB empirical study. I should first note that this does not involve the

quite distinct issue of the *burden* of proof. As CBB observe, it seems universally accepted, at least in labor arbitration (nonunion employment arbitration may vary, depending on the particular contract), that the employer bears the burden of proving there was just cause or good cause for the discharge or other discipline imposed. But on the quantum, or amount, of proof needed, there is a wide range of views.

Section 6.10 of *The Common Law*, covering quantum of proof, was hammered out after considerable discussion among the work's authors and editors. It reads essentially as follows (emphasis supplied):

(1) For *most* arbitrators, the normal quantum of proof required in disciplinary cases is "preponderance of the evidence." For a *minority*, it is "clear and convincing evidence."
(2) When the employee's alleged offense would constitute a serious breach of the law or would be viewed as moral turpitude . . . , *most* arbitrators require . . . "clear and convincing evidence." *Some* require proof "beyond a reasonable doubt" but . . . *most* hold that the criminal-law standard . . . has no place in . . . arbitration.

The Common Law adds: "Some arbitrators reject the very idea of a quantum of proof as a distracting legalism. Others argue that it is impossible to avoid at least an implicit quantum requirement."

The CBB empirical study stated that in the whole body of 2,055 discipline cases, 78 percent of the arbitrators did not specify a quantum of proof and so the authors assumed a "preponderance of the evidence" standard. Only 10 percent said "clear and convincing" and fewer than 2 percent said "beyond a reasonable doubt." That validated *The Common Law*'s conclusion as to discipline cases generally. When it came to 1,034 cases involving possible crimes or matters of moral turpitude, however, the empirical study apparently disputed the assertions of *The Common Law*. Here CBB found that 84 percent still applied the preponderance standard, while only 13 percent required clear and convincing evidence and a mere 3 percent insisted on proof beyond a reasonable doubt. Even in the 97 cases where a crime had actually been charged in the criminal justice system, the figures changed only modestly to about 76, 16, and 8 percent, respectively.

The first reaction, regrettably, has to be that my learned colleagues and I, who put *The Common Law* together, were simply

wrong, and that any future editions will have to be revised accordingly. The conclusion would thus be that the majority view is that the "preponderance of the evidence" standard prevails in all discipline and discharge cases, even those involving possible crimes or matters implicating moral turpitude. Of course, such an admission is to be made, and such a conclusion is to be accepted, only as a last resort. But seriously, I do believe that one underlying assumption in the CBB analysis must be probed before *The Common Law* throws in the towel. As the authors state it: "[W]e are assuming that an unstated quantum of proof is equivalent to a stated standard of 'preponderance of the evidence.'" That surely is not self-evident.

One of the oldest and most systematic of labor arbitration systems is the Board of Arbitration established by U.S. Steel and the Steelworkers, and chaired over the years by such luminaries as Sylvester Garrett, Al Dybeck, and now Shyam Das. The Board has strenuously resisted ever articulating exactly what is its quantum-of-proof standard. Decisions will simply say that that the Board is "convinced," "persuaded," or "satisfied," or use similar language in reaching a result. Does that mean that the same quantum is required when employees have been charged with theft, and their reputations in the community and their likely future employability are at stake, as when they are charged with excessive absenteeism? I cannot answer that question with any certainty, either for the Board of Arbitration or for all the other arbitrators who do not spell out precisely the standard that they are applying in such discipline and discharge cases. But, it really must be answered one way or another, if *The Common Law* formulation is to retain its present specificity. Perhaps a future edition will have to hedge, however, saying "some" arbitrators do this and "others" do that. In any event, I am one of those who believe that there must be some implicit standard even when none is articulated. Implicitly there could be the same "preponderance of the evidence" standard in all cases, or there could be the more demanding "clear and convincing" standard in charges of moral turpitude.

Two other factors may be at work, age and geography. I like to think that the authors and editors of *The Common Law* were a reasonably distinguished lot. But in a field like ours, distinction tends to come at a price—years in service, and gray hairs. That would be especially true of our Presidential Advisory Group. So, it is possible that some of the ideas and positions set forth in *The Common Law* reflected the thinking of a somewhat older generation than

would be true of the more heterogeneous group of 81 arbitrators who decided all the cases submitted to the Minnesota Bureau of Mediation Services since the early 1980s. Part of the trend toward the "legalization" of the arbitration process that we see as a departure from former practices unique to arbitration may include a willingness to adopt more of the standards of the civil courts. That could include heavier reliance by younger or newer arbitrators on the courts' single quantum of proof in civil cases, namely, a preponderance of the evidence, rather than the more specialized approach of veteran arbitrators in requiring a higher standard of clear and convincing proof when an employee is charged with conduct involving moral turpitude. There may well be other generational implications going well beyond quantum of proof that would be worth searching out in this empirical study.

In assessing the CBB study, another factor that may need more investigation is simply geographical, to determine whether Minnesota is a fair reflection of the country as a whole. Here I am not thinking specifically of quantum of proof—no reason immediately occurs to me why attitudes on that should vary from region to region. The Minnesota data have the advantage over published decisions that they are nearly all-encompassing for their time period and are not dependent on the subjective selection process of some editorial board. But I do have the impression that in some ways relations between employers and unions are more relaxed—dare I say "kinder, gentler"?—in the Midwest generally, and in Minnesota specifically, than in some other parts of the country.

What would be the reaction, for example, to an employer's calling the grievant as its first witness in a disciplinary case? There is certainly a logical basis for such a move—in John Kagel's colorful phrase, employers are simply trying to "nail the jelly to the tree"—to put grievants on record before they have heard the testimony of other witnesses and can adjust their stories accordingly. And, of course, there is, strictly speaking, no First Amendment right of silence in a private arbitration. But, for many persons it goes against the grain to let the employer, which has the burden of proof in a discipline case, start making its case out of the grievant's own mouth. Might arbitral attitudes on that and other matters differ from one section of the country to another? So I am not yet prepared to accept the CBB sample as necessarily universal in its application. But, at the very least, it is a goldmine of worthwhile data, and we are all much indebted to its intrepid excavators.

III. Comments on an Empirical Examination of How Labor Arbitrators Decide Discipline and Discharge Cases

Nels E. Nelson*

Laura Cooper, Mike Bognanno, and Steve Befort (CBB) have undertaken a very exciting project. They have assembled a large sample of arbitrators' decisions in discipline and discharge cases to test many of the rules we believe that arbitrators follow, many of which are discussed in *The Common Law of the Workplace*.[1] CBB will also be able to verify the results of much of the prior research on arbitral decisionmaking. What we have seen today is just a small sample of the research that is sure to follow. The Research and Education Foundation of the National Academy of Arbitrators (NAA) should be applauded for recognizing the merit of the project and providing the funding to hire several law students to do the considerable data-gathering involved in the project.

My assignment is to comment on the methodological and statistical aspects of CBB's paper. First, I will provide a few comments about their sample of arbitrators' decisions. Second, a cautionary note is offered relating to the data collection process. Finally, I will present a different view of CBB's data based on two models of the arbitral decisionmaking process.

Although CBB's sample is the largest sample of arbitrators' decisions ever assembled, three reservations about it must be noted. First, I am concerned that their sample may not be representative of all arbitrators. It consists of all of the discipline and discharge decisions in the files of the Minnesota Bureau of Mediation Services from 1982 through 2005. Members of the Bureau's arbitration panel must submit copies of all their decisions but those who are not on its panel are not required to do so. It appears that non-panel members have not submitted their decisions because the sample includes the decisions of only 81 different arbitrators over the entire 24 years included in the study. To the extent that members of the Bureau's panel are more or less experienced than

*Member, National Academy of Arbitrators, Professor of Management & Labor Relations, Cleveland State University, Cleveland, Ohio. The author wishes to acknowledge the assistance of Garima Sharma, who is a doctoral student in organizational behavior at the Weatherhead School of Management at Case Western Reserve University.

[1]St. Antoine, ed., The Common Law of the Workplace: The Views of Arbitrators, 2d ed. (BNA Books 2005).

other arbitrators or differ from non-panel members in other ways, CBB's sample is not representative of all arbitrators and may result in misleading findings.

A second question about CBB's sample arises from an examination of Table 4, where they divide their sample into three seven-year periods. The Table indicates that there were 582 cases in 1982–1989, 887 cases in 1990–1997, and 586 cases in 1998–2005. CBB offer no explanation for the 52 percent increase in the second time period or the equal decline in the number of cases in the third time period. Some of CBB's results may have been affected by whatever happened between 1990 and 1997.

A final concern about the sample is the possibility that the public sector is over-represented. Public sector cases account for a large proportion of CBB's sample and Table 6 indicates that the proportion of public sector cases rose steadily from 50 percent in 1982–1989, to 62 percent in 1990–1997, to 65 percent in 1998–2005. This may reflect what is happening in arbitration generally or it may be a function of the decisions that are submitted to the Bureau of Mediation Services. Because Table 6 finds some statistically significant differences between private sector and public sector decisions, caution may be in order in generalizing from CBB's sample to all arbitrators.

We also need to recognize the potential problems growing out of the data collection process. A team of law students analyzed the 2,055 cases included in the study, collecting data for more than 100 variables. Because important aspects of the arbitrators' decisions were not always readily apparent from reading the decisions, the students were required to do a significant amount of reading between the lines. Sometimes assumptions had to be made about what the arbitrators intended or had in mind.

CBB provided an excellent example of the problems they encountered in the data collection process. They noted that the majority of arbitrators did not state the standard of proof that they applied in their decisions. In response to this situation, CBB decided that unless an arbitrator stated otherwise, they would assume that the standard of proof the arbitrator applied was a preponderance of the evidence. To the extent that this is a bad assumption, their findings with respect to the standards of proof have to be viewed very carefully. Similar problems may exist for other variables included in the study.

CBB's paper was intended to provide only an initial glimpse of their data. The technique they chose was to compare arbitration

outcomes to selected characteristics of the arbitrators or the cases and to use chi-square statistics to test for statistical significance. For example, in Table 2 of their paper, they show the relationship between last-chance agreements and the outcomes of disciplinary grievances. The Table indicates that management won 71 percent of the cases where the grievant was on a last-chance agreement compared with only 48 percent of the cases where no last-chance agreement was in effect. Split decisions accounted for 18 percent of the cases where the grievants were on last-chance agreements and 30 percent of the cases where the grievants were not subject to last-chance agreements. The statistically significant chi-square statistic means that the outcomes of the arbitration cases were related to the existence of last-chance agreements.

Although CBB's findings with respect to the impact of last-chance agreements on arbitrators' decisions are interesting, I believe that some additional insights can be gleaned from their data by using a model that focuses more directly on the arbitral decisionmaking process rather than on arbitration outcomes. Although the usual practice in research is to develop a model as the first step in the process, I have worked backward to devise two models, one for discipline cases and one for discharge cases, to fit the data in CBB's paper.

The model of the process in a discipline case, which is shown in Figure 1, suggests that arbitral decisionmaking consists of two stages. In the first stage, arbitrators determine whether the grievants engaged in any misconduct. Where the arbitrators find no misconduct, there is no cause for discipline and the union wins the case. Alternately, if the arbitrators find that there is at least some misconduct, they may deny the grievance, upholding the discipline imposed by management, or render a split decision by reducing the penalty imposed by the employer.

In the second stage, the arbitrators determine the penalties in the cases where they decided that the grievants were guilty of at least some misconduct. One possibility is that the arbitrator will accept the penalty imposed by the employer, which means management wins the case. The other possibility is that the arbitrator will reduce the penalty meted out by management, i.e., offer a split decision.

CBB's data can easily be fitted into contingency tables based on the two-stage model. Table 1, which represents stage one, shows that where the grievants were on last-chance agreements, arbitrators found at least some misconduct in 148 cases. This corresponds

FIGURE 1. Two-Stage Model for Arbitral Decisionmaking in Discipline Cases

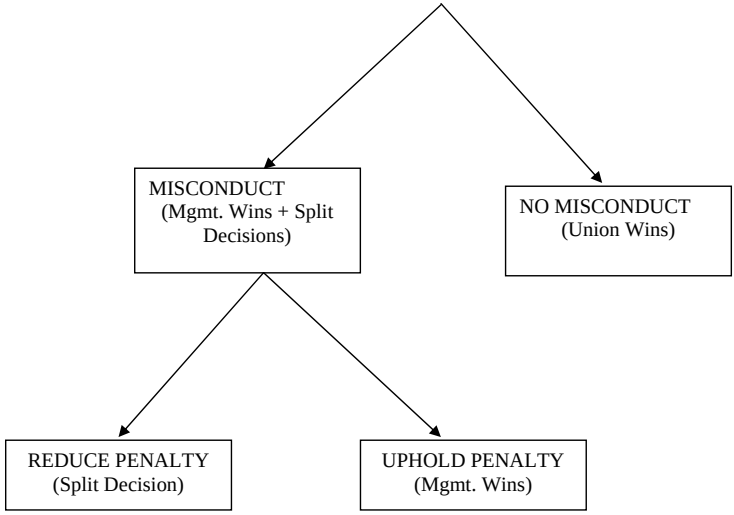


TABLE 1: Last-Chance Agreements: Stage One

	Last-Chance Agreement	No Last-Chance Agreement
Misconduct	89.1% (147)	77.6 % (1467)
No Misconduct	10.9 % (18)	22.4% (423)

chi square = 11.849, p = .001

to the 117 cases in CBB's Table 2 where management won the case plus the 30 cases where the arbitrators rendered split decisions. Arbitrators found no misconduct in the 18 cases that the union won. A similar process provides the numbers in Table 1 for grievants who were not on last-chance agreements.

Table 1 suggests that last-chance agreements influence arbitrators' findings regarding misconduct. It reveals that in 89.2 percent of the cases where the grievants were on last-chance agreements, the arbitrators found that there was at least some degree of misconduct by the grievants, while in cases where the grievants were not on last-chance agreements, only 77.5% percent of the arbitrators concluded that there was misconduct. The most important number in the table is the chi-square statistic. It is significant at the .000 level, which means that we can safely conclude that arbitrators were more likely to find misconduct where the grievants were on last-chance agreements than when they were not.

The results for stage one are surprising. The issue of whether there is misconduct on the part of the grievant is a factual determination based on the testimony and evidence offered by the parties. Prior misconduct should not be expected to influence this determination. Table 1, however, suggests that arbitrators are more likely to find grievants guilty of misconduct when they have been involved in prior misconduct as reflected in their being subject to last-chance agreements.

Table 2 represents the second stage of the decisionmaking process. It indicates that when grievants were on last-chance agreements, there was no reduction in the penalty imposed by the employer in 117 cases. This corresponds to the cases the employer won in CBB's Table 2. The 31 cases where the penalty was reduced are the split decisions in CBB's table. A similar process provides the data for the grievants who were not on last-chance agreements at the time of their discipline.

Table 2: Last-Chance Agreements: Stage Two

	Last-Chance Agreement	**No Last-Chance Agreement**
No Reduction in Penalty	79.6% (117)	61.7% (905)
Reduction in Penalty	20.4% (30)	38.3% (562)

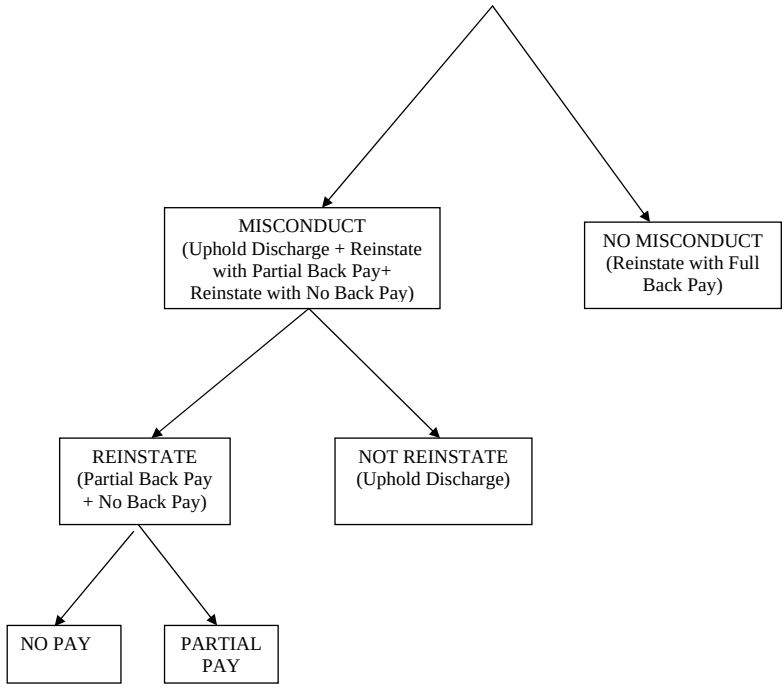
chi square = 18.435, p = .000

The results for stage two are shown in Table 2. It indicates that grievants on last-chance agreements had their penalties reduced in 20.4 percent of the cases, while those not on last-chance agreements had their penalties reduced in 38.3 percent of the cases. Again, the important number is the chi-square statistic, which is significant at the .000 level. Thus, the penalties imposed by the arbitrators were related to the existence of last-chance agreements.

The stage two results are as expected. Arbitrators universally regard a grievant's prior disciplinary record as an important factor in assessing the propriety of a penalty imposed by an employer. Because a last-chance agreement represents prior misconduct by a grievant, it is not surprising that arbitrators less frequently reduce the penalties imposed by employers when grievants are subject to last-chance agreements.

CBB also considered the impact of membership in the NAA on arbitrators' decisions in the 1,432 discharge cases in their sample. The analysis of a discharge case requires a three-stage model, which is shown in Figure 2. Stage one is essentially the same as the

FIGURE 2. Three-Stage Model for Arbitral Decision Making in Discharge Cases



two-stage model for all discharge cases. If there is no misconduct, then the union wins the case. At stage two, the question is whether or not a grievant who is guilty of misconduct is reinstated. A grievant is not reinstated when the employer wins the case. Stage three simply involves the resolution of the back pay issue for those who are reinstated. Some receive partial back pay and others receive no back pay.

In Tables 3, 4, and 5, the results from CBB's Table 10 are placed in the contingency tables corresponding to each stage of the discharge model. Table 3 indicates that at stage one, 82.8 percent of the NAA members found the grievants guilty of misconduct compared with 76.9 percent of the nonmembers. This is statistically significant, suggesting that NAA members are a tough bunch!

In stage two, which is shown in Table 4, NAA members reinstated the grievants 34.4 percent of the time, while non-members

Table 3: NAA Membership: Stage One

	NAA Members	Non-Members
No Misconduct	17.2% (132)	23.1% (151)
Misconcduct	82.8% (637)	76.9% (503)

chi square = 7.78, p= .005

Table 4: NAA Membership: Stage Two

	NAA Members	Non-Members
Reinstate	34.4% (219)	33.8% (170)
Not Reinstate	65.6% (418)	66.2% (333)

chi square= .042, p= .837

reinstated the grievants 33.8 percent of the time. The difference is not statistically significant, indicating that there is no relationship between NAA membership and the decision to reinstate the grievants, i.e., members are no more or less likely to reinstate the grievants than nonmembers. This suggests that we are not so bad after all!

In stage three, shown in Table 5, the issue is back pay for those who are reinstated. At this stage, 71.7 percent of the members compared with 48.8 percent of the nonmembers reinstated the grievants without back pay. The difference is statistically significant, indicating that members are less likely to grant back pay when they reinstate the grievants. Are we tough or are we trying to make everyone happy with a compromise?

Although the results with respect to NAA membership are interesting, they suffer from a common statistical problem. The implicit assumption involved in the simple statistical tests that CBB and I have used, as well as the more sophisticated work done by other researchers, is that arbitrators are assigned to cases randomly. Because arbitrators are specifically chosen to hear certain types of cases, we have a problem that the statisticians refer to as "sample selection bias." There are sophisticated techniques that address this situation, but they go far beyond the scope of this comment.

CBB also attempted to sort out the conflicting results of the prior research regarding the impact of a grievant's gender on an arbitrator's decision in a discharge case. As they noted, some researchers found that arbitrators were more likely to reinstate females than males, while other researchers concluded that gender had

TABLE 5: NAA Membership: Stage Three

	NAA Members	Non-Members
Partial Pay	28.3% (62)	51.2% (87)
No Back Pay	71.7% (157)	48.8% (83)

chi square= 21.175, p= .000

no impact on arbitrators' decisions.[2] Although the results shown in Table 7 of CBB's paper are close to the 5 percent probability level, which is generally required to conclude that the results are not due to chance, the appropriate conclusion is that there is no relationship between the sex of a grievant and the outcome of the arbitration process. When CBB's data are plugged into the three-stage model for discharge cases, the results indicate that gender is not related to any stage of the arbitral decisionmaking process.

The fact of the matter is that the results relating to the impact of a grievant's gender on an arbitrator's decision illustrates a significant shortcoming of the methods used in CBB's paper and in this comment. As CBB indicate, we cannot ignore factors other than gender that may influence arbitrators' decisions; rather, we need to control for those factors in order to estimate the independent impact of gender on arbitrators' decisions. For example, if arbitrators appear to treat females differently, it may be that the different outcomes are a result of the offenses they commit or differences in the types of jobs they hold rather than their gender. There are multivariate statistical techniques that address this problem and CBB are already working on a paper using much more sophisticated statistical techniques.

Despite these modest reservations, it is clear that CCB are to be commended. They have assembled an astounding sample of arbitrators' decisions. In their present paper, we have seen only the tip of the iceberg. CBB will be busy for many years attempting to answer questions that have never been answered and, in many cases, have never been asked. We will also be able to see if the things we have always believed about arbitrators and their decisions are true.

---

[2]Studies by Bemmels, *The Effect of Grievants' Gender on Arbitrators' Decisions*, 41 *Indus. & Lab. Rel. Rev.* 251 (1988), and Ponak, *Discharge Arbitration and Reinstatement in the Province of Alberta*, 42 *Arb. J.* 39 (June 1987), found that arbitrators treat females more leniently. Block & Stieber, *The Impact of Attorneys and Arbitrators on Arbitration Awards*, 40 *Indus. & Lab. Rel. Rev.* 543 (1987); Scott & Shadoan, *The Effect of Gender on Arbitration Decisions*, 10 *J. Lab. Res.* 429 (1989); Nelson & Uddin, *The Impact of Delay on Arbitrators' Decisions in Discharge Cases*, 23 *Lab. Stud. J.* 3 (1998); and Zirkel & Breslin, *Correlates of Grievance Arbitration Awards, J. Collective Negations* 45 (1995), found no relationship between grievants' gender and arbitration outcomes.