New Zealand Journal of Industrial Relations, **22(3)/23(1)**: 277-300

Gender Patterns in the New Zealand Employment Tribunal: Some Notes on Theory and Research

Ian McAndrew, Dermott J. Dowling and Sean Woodward *

Introduction

The New Zealand Employment Tribunal has now been in operation for over six years since its inception, along with the Employment Court, under the Employment Contracts Act 1991 (the ECA or the Act). Other than basic record keeping by the Tribunal, and by the Industrial Relations Service of the Department of Labour, there has been little systematic analysis or evaluation of the output of the Tribunal's adjudication juris-diction. As the superior body that interprets the law, there has been more study of the Employment Court, at least of the legal substance of its decisions, but little of the Tribunal in its application of the law.

Against this sparse background, Morris (1996) conducted a multi-faceted investigation of gender bias in the Court and the Tribunal. A necessarily short summary of Morris' paper will no doubt do it less than full justice, but it is an important prelude to the present note in which we seek to draw attention to relevant overseas work on gender influence in employment arbitration or adjudication, and to introduce some further New Zealand data on the subject.

Morris examined the limited number of published Court and Tribunal personal grievance cases for the years 1991 through 1994, together with unpublished cases with an obvious gender factor – those alleging discrimination on the basis of gender and those alleging sexual harassment. She found first that applicant gender did not appear to influence the chances of having a grievance sustained by the Court or Tribunal in a judicial forum. Second, there was no apparent gender factor associated with whether a manager's decision, challenged through the personal grievance procedure, was upheld or overturned by the Court or Tribunal. Third, figures on reimbursement to successful grievants of monies lost as a consequence of their grievances, a calculation that is in part formulaic and in part discretionary, showed no evidence of gender bias. Fourth, Morris concluded that women appeared to be awarded reinstatement less often than were men, although she recognized that the nine reinstatements looked at constituted a very small sample.

* The authors have all been affiliated with the Department of Management, University of Otago. Dermott J. Dowling and Sean Woodward were research associates. Dowling is now with the New Zealand Dairy Board. Woodward is with the Wellington law office of Chapman Tripp. Ian McAndrew is a Senior Lecturer in Industrial Relations, and is also a Member of the Employment Tribunal. The views expressed are the personal views of the authors. The Foundation for Industrial Relations Research and Education (NZ) has provided funding for the on-going research reported in this note. The authors are grateful to Valerie Thompson for statistical analysis and valuable insights.

Morris next considered compensation awards to successful grievants under two heads – for loss of benefits, and for humiliation, loss of dignity and injury to feelings – with the focus on the latter as the more discretionary award. Morris appraised the compensation figures as "present(ing) the impression that this area is free from gender bias" (Morris, 1996: 75), but then moved on to consider "what the statistics do not reveal."

Morris concluded that the "unstated system" of factoring grievant income into the calculation of compensation for humiliation, loss of dignity and injury to feelings (as against compensation in the more tangible loss of benefits category) inherently and inappropriately disadvantages women grievants. The paper proceeded then to evaluate selected comments of decision makers, comparing awards and comments from one case to another, and finding in this material evidence of "a distinct pattern in the employment institutions of valuing the hurt experienced by women as being worth less than the hurt experienced by men" (Morris, 1996: 78). In the balance of the paper, Morris examined the small numbers of gender discrimination and sexual harassment cases and generally concluded, following Davis (1994), that the institutions have not handled these cases well, but she added that their performance in these regards may be improving.

Scrutiny of the employment institutions, including for gender bias, is an entirely appropriate, indeed important, function for academic and policy research. Neither the examination nor findings need imply criticism of either decision makers or their institutions. The social value of research of this nature as feedback to the judges and members of the institutions relies in very large part on its credibility, which in turn relies on the quality of its methodology, its data, and its conclusions reasonably drawn. The Morris investigation is most valuable, in our view, for having initiated (at least in New Zealand) a systematic examination of gender bias in what might be referred to as the "general product" of the institutions' adjudicial function, namely the great majority of personal grievance and other individual rights cases in which gender is not nominally a factor, and ought not to be a factor at all.

While new to New Zealand, examination of gender bias in employment or labour arbitration has a substantial history in academic research, particularly in North America. Historically, most research on arbitration focused on the relationship between characteristics of the arbitrator – age, affiliations, education, experience and so on – and his or her (but usually his) decisions. Research attention in the 1970s and early 1980s turned to the decisions of arbitrators in high gender profile cases, most obviously gender discrimination and sexual harassment cases, reflecting the increasing awareness and documentation of discrimination against women across a broad array of employment related decisions.

As Bigoness and DuBose (1985) point out, by the mid-1980s a more systematic and sophisticated approach was emerging in which researchers sought to examine whether grievant gender was related to or influential in arbitration decisions over the normal range of decisions, principally unjustified dismissal, that women workers share with men. And Bigoness and DuBose, following Zirkel (1983) introduced another dimension to their research: a consideration of whether any decision patterns emerged by arbitrator gender.

Given that the vast majority of arbitrators in North America to that point had been men, arbitrator gender had not attracted a lot of research attention.

The point of this present paper is to examine some hard data for gender patterns – incorporating both applicant and adjudicator gender – in the general product of the Employment Tribunal's adjudication jurisdiction. Given the critical nature of the subject, we believe it important that scrutiny of the Tribunal ought to be properly informed by the body of available theory and research, and we preliminarily review that material in some depth.

The theoretical framework

The theory relevant to gender bias in labour or employment arbitration is derived in very large part from the administration of criminal justice literature. While some of the research into gender bias in arbitration has been atheoretical (Bemmels, 1990), much of it has been driven by two primary competing theories (Bemmels, 1988a).

The first such theory is termed *the chivalry/paternalism thesis*. Its essence is the perception of men as protectors of women, the latter being seen as weaker and not entirely accountable for their actions. The implication of the chivalry/paternalism thesis is that male arbitrators will be more tolerant of offenses committed by women grievants than of offenses committed by men. While it might be argued (Moulds, 1980) that chivalry and paternalism are different concepts, only the latter being clearly negative in its necessary power relationship, the predicted effect in arbitration is the same. Under this thesis, male arbitrators would be expected to be more lenient towards women grievants than men grievants, in other words to discriminate in favour of women grievants.

The competing thesis from the administration of criminal justice literature is termed *the evil woman thesis*. The implication of this thesis is the opposite of the other: women grievants will be treated more harshly than men grievants (by male arbitrators), particularly in terms of punishment once having been determined to be guilty of the offence alleged by the employer. The essence of the evil woman thesis is that women are not expected to commit offenses, and that women who do so violate this stereotypic expectation about the proper behaviour of women. When a woman does commit an offence, she is no longer given the preferential treatment normally extended to women. To the contrary, she is penalized not only for the offence, but for her inappropriate gender role behaviour as well, resulting in harsher punishment than is given to men grievants found to be guilty of similar wrongdoing. Male arbitrators, then, will discriminate against women grievants, at least in sentencing decisions once becoming convinced that the grievant is guilty of wrongdoing.

As noted, these competing theses are both based on stereotypic assumptions about the sexes, and they have been researched and presented as relevant to male judges and arbitrators only. Oswald and VanMatre (1990) introduced into arbitration research a third theory that hypothesised possible gender bias in female arbitrators: *the queen bee thesis*.

This thesis contends that women in authority positions place extremely high expectations on other women because they themselves worked hard to achieve their positions of rank and authority. According to this theory, a successful woman is usually in the best position to advance the cause of less powerful women, but is least inclined to do so. However, when a woman is perceived as already successful, the "queen bee" is likely to believe that she has earned her success and will support her. Oswald and VanMatre translated the queen bee thesis as hypothesising that female arbitrators would be expected to hold higher standards for women grievants than for men grievants, resulting in harsher treatment for women grievants. In other words, this thesis predicts that female arbitrators will discriminate against women grievants.

The research background

The available research on gender bias in employment arbitration is almost entirely North American, and much of it has been organised around the theses outlined above. Its centrepiece is a series of research projects by Bemmels published between 1988 and 1991, and the full body of research can be presented sequentially in three parts: papers preceding the Bemmels projects, the Bemmels projects themselves, and work that appeared contemporaneously with the Bemmels papers or subsequently. For the most part North American arbitrators operate in private practice and are engaged privately by the parties to a dispute or grievance. They have operated primarily in the unionised sector, though that is changing, and they are guided by an extensive case law of arbitration decisions, though in recent years employment protection legislation has increasingly emerged as the unionised sector has reduced, particularly in the United States. Private arbitrators in North America are, in most respects that are relevant to our research, equivalent to Employment Tribunal adjudicators in the New Zealand setting.

The pre-Bemmels studies

Rosenberg (1979), himself an arbitrator, first asserted on the basis of a small sample and his own experience that (male) arbitrators discriminate against women grievants. Rodgers and Helburn (1984) gave some support to that assertion. Their sample was the 37 dismissal appeals from five unionised refineries in the American southwest over the period 1975 to 1981. The dependent or "outcome" variable was the single dichotomous one: was the employee reinstated or not. Rodgers and Helburn concluded that female grievants were less likely to be reinstated than were male grievants, a conclusion that was consistent with the observations of company and union officials they interviewed at the refineries.

Ponak (1987) came to a contrary conclusion. His sample consisted of all recorded dismissal arbitration awards in the Canadian province of Alberta between 1982 and 1984. There were a total of 159 grievants involved, of whom about one-third were women. Ponak used a trichotomous outcome variable: either the dismissal was upheld, or the employee was reinstated with full reimbursement of monies and privileges lost, or the employee was

reinstated but with less than full reimbursement. This last option, not uncommon in North American arbitrations, amounts to the arbitrator substituting a lesser form of discipline, essentially a disciplinary suspension, in place of the dismissal imposed by management. Ponak tested several independent variables, both characteristics of the grievant and contextual factors, and concluded that gender was the only factor related to decision outcome in a statistically significant manner. He found that male grievants were more than twice as likely as female grievants to have their dismissals upheld. While not specified, it can probably be accepted that the vast majority of the arbitrators in Ponak's early 1980s study were men.

Meanwhile, Zirkel (1983) pioneered the study of decision-maker gender in labour or employment arbitration. Zirkel built profiles of 400 grievance arbitration cases, noting characteristics of the hearing, the issues involved, decision outcomes, and the gender, occupation and experience of the arbitrator. The case sample involved 225 arbitrators, about seven percent of whom were women. Zirkel concluded that the gender of the arbitrator did not appear to be related to outcomes.

Bigoness and DuBose (1985) used students in a laboratory experiment to examine the effects of both grievant gender and arbitrator gender on arbitration outcomes. They drew three specific hypotheses for testing: first, that arbitrators would treat women grievants more leniently than men grievants; second, that women arbitrators would be more lenient than men arbitrators; and third, that arbitrators of each gender would treat same gender grievants more leniently than other gender grievants. In the event, they found no support for any of the hypotheses, or indeed for any gender effect, save that women students serving as arbitrators tended to see the case offence (drinking alcohol on the job) as less serious than did their male counterparts.

Suffice it to say that these first phase research results on gender bias in employment arbitration were less than unanimous. Next came Bemmels, and the results of his several somewhat similar studies showed more consistent results. By and large, Bemmels organised his projects around hypotheses drawn from the chivalry/paternalism thesis, and he endeavoured to control for a range of variables in order to isolate the effects of grievant gender. In two of his studies, the gender of the arbitrator was also incorporated as a variable.

The Bemmels studies

Bemmels (1988a), relying on the chivalry/paternalism thesis, hypothesized that (male) arbitrators in dismissal cases would, first, be more likely to reinstate women grievants than men; and second, levy lesser penalties against women than men where reinstatement with less than full reimbursement was awarded. His data was all 126 arbitration decisions in dismissal cases filed with appropriate authorities in Alberta between the beginning of 1981 and mid-1983. In all cases, the arbitrator or the chairperson of the arbitration panel was a man. Bemmels employed three outcome measures: first, whether the grievance was

sustained or denied; second, where reinstatement was ordered, whether it was full reinstatement or partial reinstatement (that is, reinstatement but with a disciplinary suspension); and third, where reinstatement was only partial, the length of the disciplinary suspension imposed by the arbitrator. Bemmels controlled for several variables: the reason for the discharge, the grievant's occupation, whether the employment was in the public or private sector; whether the case was heard by a single arbitrator or an arbitration panel; and the year of the decision. Bemmels (1988a: 259) concluded as follows:

> The empirical results suggest that women received more favorable treatment by arbitrators on all three dependent variables analyzed. When the other variables are controlled for, women were twice as likely as men to have their grievances sustained and 2.7 times more likely to receive a full reinstatement, and in cases in which a suspension was imposed by the arbitrators, women received, on average, a suspension 2.1 months shorter than that for men. These results contradict the simulation results of Bigoness and DuBose (1985), who found virtually identical outcomes for male and female grievants. The results are consistent, however, with the chivalry/paternalism thesis discussed above, and also with the empirical results of studies investigating gender effects on the decisions of judges.

Bemmels recognized that his conclusion of bias in favour of female grievants assumed a number of things, most obviously that the men and women grievants had, on average, equally valid cases. And he further acknowledged that a number of variables not controlled for, including the age, seniority and record of the grievant, or the relevant rules or contractual provisions behind the dismissals, might well have been influential in the decision outcomes. Nonetheless, these first findings provided strong support for the thesis that male arbitrators discriminate, not against women grievants, but in their favour.

Bemmels (1988b) tested essentially the same hypotheses using a sample of published discharge arbitration awards in the United States dated between 1976 and 1986. Bemmels also hypothesized that a diminishment of gender bias would be apparent over time. The sample included 1,812 cases decided by 729 arbitrators. Of the arbitrators, 45 were women and they decided 92 of the cases in the sample. Of the grievants, 312 were women. Of the possible gender combinations, the male grievant and male arbitrator combination accounted for almost 80 percent of the cases. Bemmels controlled for the same variables as he had in the previous study, and also included the grievant's disciplinary record, the industry group of the employer, and a dummy variable for mitigating factors acknowledged by the arbitrator. These last typically included inconsistent enforcement by the employer, circumstances negating intent, provocation, long service with the employer, and a history of good performance. The outcome variables employed were the same as those in the earlier study.

The conclusions from Bemmels (1988b) in important respects mirrored those from the previous study. He found that women grievants were more likely than men to have their grievances sustained by male arbitrators, but no more likely than men to have their grievances sustained by female arbitrators. In cases in which the grievance was sustained, women grievants were more likely than men to be awarded a full reinstatement by male arbitrators, but again no more likely than men to be awarded a full reinstatement by female arbitrators. There were, at the same time, some findings that were contrary to Bemmels'

expectations. In the early years, women grievants who were awarded only partial reinstatement in fact received longer suspensions than men from male arbitrators, although later in the period there was no gender pattern in this respect. Though Bemmels did not suggest it, a causal link between the more likely reinstatement of women and the longer suspensions imposed on women in substitution for their dismissals seems plausible, and consistent with the more favourable treatment of women by male arbitrators on the other outcome dimensions. On the longitudinal dimension, only this more lenient treatment of women grievants on suspensions showed up as a trend, and that was also in a direction contrary to that hypothesized.

Bemmels (1988c) tested essentially the same hypotheses using reported discipline arbitration decisions, some involving dismissals and others involving disciplinary measures less than dismissal, for the period 1977 through 1982 in the Canadian province of British Columbia. The sample consisted of 633 decisions, all issued by male arbitrators or arbitration boards of which the chairperson was a man. The outcome variables employed were essentially those used in the previous studies. Variables controlled for were similar to those in Bemmels (1988b), but with an additional variable being the original discipline imposed by the employer. In this study, Bemmels found that the odds of women grievants receiving a full exoneration rather than having a lesser discipline substituted for the discipline originally imposed by the employer were 80 percent higher than the odds for men. In all other respects, however, there were no gender effects apparent this time.

Bemmels (1990) was a more general expeditionary survey of the relationships between various arbitrator characteristics and arbitrator decisions. It was seen as atheoretical and so no hypotheses were proposed. The case data source was essentially that used in Bemmels (1988b), reported United States cases from 1976 through 1986, but here the sample was extended to include disciplinary suspension cases in addition to discharge cases. Two thousand and one cases decided by 459 arbitrators constituted the sample. Decision outcomes were analyzed against a wide range of arbitrator characteristics, 27 in all including gender, with the usual range of variables controlled for, including the grievant's gender. Bemmels reported a number of statistically significant relationships between arbitrator characteristics and their decisions, including that women arbitrators gave suspensions almost nine weeks shorter than men arbitrators did. Putting things in perspective, however, the arbitrator characteristics variables collectively, including gender, explained only about five percent of the variation in the outcome variables.

Finally, Bemmels (1991) tested the same hypotheses tested in Bemmels (1988b), including the expectation that male arbitrators' more favourable treatment of women grievants would be seen to diminish over time. This time the sample was 557 disciplinary suspension cases, again drawn from the published United States arbitration reports for the years 1976 through 1986. The cases were decided by 322 arbitrators, of whom 16 were women. The outcome variables were the same as in the previous studies, allowing for the fact that this sample dealt with suspension cases rather than dismissals. The usual control variables were covered. In the only statistically significant result, Bemmels reported that male arbitrators were more likely to sustain the grievances of women than of men, but there was no

difference in the treatment of men and women grievants by female arbitrators. But this more favourable treatment of women grievants by male arbitrators did not extend to other outcome measures: full versus partial exoneration or the length of lesser suspensions substituted by the arbitrator. And no change was apparent over time.

While Bemmels results were not uniform across his several studies, the overwhelming impression from his body of work on the subject was of more favourable treatment of women grievants than of men grievants by male arbitrators. Though it was less of a focus and there was less data to work with, no grievant gender patterns were found for female arbitrators. By extrapolation, Bemmels could be said to have identified differences by grievant gender between male and female arbitrators: he generally found that male arbitrators discriminate in favour of women grievants while female arbitrators do not. It should be noted that Caudill and Oswald (1992) challenged Bemmels' methodology, but confirmed the thrust of his findings. Other research contemporaneous with or following Bemmels has produced mixed results, but cannot be said to have resoundingly supported his findings.

The post-Bemmels studies

Scott and Shadoan (1989) sought to examine the effects of both grievant gender and arbitrator gender on arbitration decisions using 169 published United States arbitration awards. Their objective was to use 50 cases in each grievant-arbitrator gender combination but there were insufficient woman grievant-woman arbitrator cases in their sources. Citing theory and previous research as discussed above, Scott and Shadoan hypothesized that arbitrators would treat women grievants less severely than men grievants, that women arbitrators would be more lenient than men arbitrators, and that arbitrators of each gender would treat same gender grievants more favourably than other gender grievants. They found no significant results supporting any of these hypotheses or any other gender effects on the arbitrators' decisions.

Thornton and Zirkel (1990) surveyed the decisions of arbitrators, and the predictions of managers and union officials as to the likely arbitral decisions on three variants of two grievance scenarios: a dismissal for absenteeism and a contract interpretation dispute involving promotion. The authors reported no relationships between arbitrator gender and their decisions.

Oswald and VanMatre (1990) set out to test the queen bee thesis by soliciting decisions from female arbitrators on a hypothetical drug-testing dismissal case with grievant gender being the independent variable. Half of the sample of 103 female arbitrators received the case with a female grievant; the other half with a male grievant. The outcome variables were the usual ones familiar from the Bemmels studies. Twenty nine arbitrators responded with decisions. Oswald and VanMatre found to a statistically significant level that the women arbitrators were more likely to award a full reinstatement to a woman grievant than to a man, once having determined that the dismissal should be overturned. On the other

outcome variables – whether to sustain or overturn the dismissal, and the length of suspension where the decision is for reinstatement with less than full reimbursement – there were no patterns by grievant gender.

Caudill and Oswald (1993) was an extension of the study reported in Oswald and VanMatre (1990). In addition to the 103 female arbitrators, 400 male arbitrators were also sent the drug-testing case; again for one half the grievant was a man and for the other half the grievant was a woman. A total of 146 arbitrators responded with decisions, the 29 women and 117 men. Respondents also completed a biographical questionnaire, and various arbitrator characteristics were built into the model along with grievant gender. The authors concluded in relation to gender that women grievants are more likely than men to have their dismissals overturned by the total population of arbitrators (but not, according to the previous paper, by the sub-population of women arbitrators only), and that women arbitrators are less likely to fully reinstate grievants than are their male counterparts.

Crow and Logan (1994) conducted a multi-faceted study that looked principally for any links between various arbitrator characteristics including gender and prior decision record, type of representation engaged by the parties for the hearing, and grievant gender and arbitral outcomes. On the gender dimension, citing the chivalry/paternalism thesis and Bemmels' work, the authors hypothesized that male arbitrators would be more likely to sustain the grievances of women grievants than of men. Their sample was 248 published United States arbitration awards in discipline cases where alcohol or drug use was a basis for discipline. They found no gender effects at all.

Finally, Steen, Perrewe and Hockwarter (1994) examined 603 United States arbitration awards in discipline cases published in the five year period to mid-1992. Of the 603 grievants, 111 were women. Steen et. al. looked at three principal variables: whether the discipline being appealed was a dismissal or something less than dismissal, the outcome in terms of whether the grievance was upheld, denied or modified, and the gender of the grievant. The authors reported no gender effects.

It should be noted that there is a considerable peripheral literature that has not been examined here: related research on gender bias at various points in the justice system, particularly the criminal justice system; research on gender bias in a range of employment and personnel management decision-making; and empirical studies on the relationships between arbitration outcomes and grievant and arbitrator characteristics beyond gender. The point here has been to review those studies expressly on the point of gender patterns in employment or labour arbitration. The results of this research would have to be said to be mixed.

The earliest papers (Rosenberg, 1979; Rodgers and Helburn, 1984) tended to suggest discrimination against female grievants by male arbitrators, but there is no further evidence of that in the subsequent research. Rather, the three predominant themes that emerge are: first, discrimination in varying degrees and ways in favour of female grievants by male arbitrators (Ponak, 1987; Bemmels 1988a, 1988b, 1988c, 1991; and Caudill and Oswald,

1993); second and contrary, the finding of no effects of grievant gender at all (Bigoness and DuBose, 1985; Scott and Shadoan, 1989; Crow and Logan, 1994; and Steen, Perrewe and Hockwarter, 1994); and third, with exceptions, the finding of no relationship between arbitrator gender and arbitration outcomes (Zirkel, 1983; Bigoness and DuBose, 1985; Scott and Shadoan, 1989; Thornton and Zirkel, 1990; and Crow and Logan, 1994). These, it seems to us, are the themes that ought principally to guide research on possible gender bias in the adjudication function in the New Zealand employment institutions. We turn next to presenting some early data on point.

Some New Zealand data

The material presented below is drawn from a database of Employment Tribunal adjudication decisions under continuous construction at the Department of Management, University of Otago with sponsorship from the Foundation for Industrial Relations Research and Education (NZ). In brief summary, the intent for the database is to construct profiles for all Tribunal adjudication decisions on a wide range of variables representing characteristics of case issues, parties, hearings, and outcomes. The data are drawn from case summaries prepared and published by the Employment Institutions Information Center and from the full-text decisions. The data are kept in SPSS format for analysis.

The sample

For this analysis we have extracted 530 relatively early decisions from the database. These are all of the substantive decisions (as against costs or preliminary or procedural matters) issued by the Tribunal during calendar years 1993 and 1994 on individual unjustified dismissal grievances, where the primary stated reason for dismissal was clearly identifiable as either misconduct-based, performance-based or redundancy-based, and in which the gender of the grievant was unequivocally determinable. Limiting the study to dismissal grievances has the advantage of providing some initial broad controls. Applicable rules and contract language, and the methods of case analysis and writing tend to be broadly similar in most dismissal cases. Limiting the sample to 1993 and 1994 limits any longitudinal effects, such as those hypothesised by Bemmels in some of his work, although we would intend to examine that angle in future work.

Outcome variables

We recognize that gender bias may take forms and have consequences not directly measurable in quantitative terms, and that qualitative analysis of adjudicatorial comments of the type undertaken by Morris (1996) may in some respects be appropriate in searching for gender bias. Such qualitative analysis is well beyond our scope and sophistication. We focus instead on the hard data that is available to us. We have employed two primary outcome variables for this report. The first is the simple dichotomous one: was the

grievance sustained in whole or part or was it denied; or in other words, did the employee win or lose. The employee lost in 154 (29 percent)[1] of cases in our sample, and won to some or all extents in the remaining 376 cases (71 percent).

Our second outcome variable is the net level of compensation awarded to successful grievants for humiliation, loss of dignity and injury to feelings under section 40 (1) (c) (i) of the ECA. We agree with Morris (1996) that, as the most discretionary category for the award of monies to successful grievants, this head is the most appropriate one to test for gender patterns. The distribution of compensation awards among the 376 successful grievants is set out in Table One.

Table One: Section 40(1)(c)(i) compensation awarded

Amount of compensation	Number of Cases
No compensation	40
$1 – $5,000	222
$5,001 – $10,000	83
$10,001 & above	31

Reinstatement was awarded in only 17 of the cases in our sample and, with so few decisions to work with, we have elected not to include reinstatement as a measured outcome for purposes of this paper. Again, it is obviously an appropriate subject for future work.

Independent variables

We have examined and attempted to control for a range of independent variables in an effort to isolate any patterns by the focal variables: grievant or employee gender and adjudicator gender. The independent variables that we have examined, other than the gender variables, are: stated reason for dismissal, geographic Tribunal jurisdiction, occupation of the dismissed employee, and nature of representation at hearing. Intuitively, these seem to us to be measured variables that might have a bearing on case outcomes. We describe the sample first in its gender dimensions, and then in terms of those other variables.

---

[1] We have rounded all percentages in the text to the nearest half percent; percentages in tables remain precise.

Of the 530 decisions in the sample, 434 (82 percent) were issued by male adjudicator members of the Employment Tribunal and 96 (18 percent) were issued by female members.

That would be about proportional to the average gender makeup of the Tribunal during the 24 month period of calendar 1993 and 1994.

Of the 530 cases in the sample, 357 (67.5 percent) were brought by male grievants and 173 (32.5 percent) were brought by female grievants. That is an "underepresentation" of women relative to workforce makeup, though far less of an underepresentation than was typical in most of the North American studies reviewed above. Some of the North American literature cited a possible union bias against would-be women grievants as one factor in the underepresentation of women amongst grievants. That ought not to be a factor under the individual representation regime of the ECA. There may still be several explanations. Women may commit fewer offences warranting dismissal. Or they may opt for mediation in greater numbers than male grievants. Certainly there is literature supporting both of these explanations. And there are probably other explanations available as well.

Turning to the other independent variables, 255 of the 530 cases in our sample (48 percent) were for alleged misconduct of one sort or another, 142 (27 percent) were for alleged poor performance, and 133 (25 percent) cited redundancy as the reason for dismissal.

The Auckland Tribunal issued 253 of the decisions in the sample (48 percent), 119 (22.5 percent) were out of the Wellington Tribunal, and the remaining 158 (30 percent) were issued by the Christchurch Tribunal (Christchurch and Dunedin offices). It should perhaps be noted that there were no female Tribunal adjudicators based in the Christchurch Tribunal during the period under study.

The occupation of the applicant was not clearly identifiable in 15 cases (3 percent of the sample). For the rest, we used a four-category system. We classified 122 applicants (23 percent) as "managers and professionals", 83 (15.5 percent) as "supervisory and white-collar", 141 (26.5 percent) as "pink collar" (services and sales), and 169 (32 percent) as "blue collar" (manufacturing, trades, equipment operators, agricultural workers).

The final independent variable that we have tried to control is the parties' representation at the adjudication hearing. While we will examine this variable more fully in the future, we have kept it at the simple dichotomous level for purposes of this analysis, tracking only two categories: independent professional advocates with legal qualifications and independent professional advocates without legal training. Other representation options, including self-representation or in-house representation, were not specifically addressed for this analysis. Of the 530 applicants in our sample, 235 (44.5 percent) were represented at hearing by lawyers, while 199 (37.5 percent) were represented by advocates who were not lawyers. The rest (96 or 18 percent) chose a different representation option. Of the 530 respondents, 272 (51.5 percent) were represented by lawyers, 130 (24.5 percent) were represented by non-lawyers, and 128 (24 percent) chose other options.

While we have incorporated these several variables into our analysis, it goes almost without saying that we are nonetheless necessarily making some very substantial assumptions about the sample in presenting and analysing the data that follows. Most obviously, we are assuming that applicants have equally meritorious cases across both applicant and adjudicator genders, as well as secondarily across reasons for dismissal, geographical jurisdictions, occupations, and representation options.

Hypotheses for testing

We drew three hypotheses to test the themes that showed through in the theoretical and research literature reviewed above:

> *Hypothesis One:* Male adjudicators treat female grievants more favourably than male grievants on adjudication outcomes.
>
> *Hypothesis Two:* Female adjudicators treat male grievants more favourably than female grievants on adjudication outcomes.
>
> *Hypothesis Three:* Female adjudicators treat grievants more favourably than do male adjudicators on adjudication outcomes.

There is no theoretical base to Hypothesis Three. The weight of the research on point suggested no relationship between arbitral outcomes and arbitrator gender. For testing purposes, we have simply chosen to state the hypothesis as an affirmative.

Initial gender patterns

We initially examined the data for any correlations between the gender variables and the outcome variables. The frequencies for win-lose outcomes and applicant and adjudicator gender are set out in Tables Two and Three. Each of these relationships proved statistically significant in correlation analysis, with Chi-Squares of $<.005$ for applicant gender and $<.05$ for adjudicator gender. So, female applicants in our sample were significantly more likely than male applicants to have had their grievances upheld, while female adjudicators were somewhat less likely than male adjudicators to have upheld grievances.

Table Two: Win-Lose Outcomes by Applicant Gender

		Applicant Gender	
		Female	Male
Employee Won	Count Percent	138 (79.8)	238 (66.7)
Employee Lost	Count Percent	35 (20.2)	119 (33.3)

Table Three: Win-Lose Outcomes by Adjudicator Gender

		Adjudicator Gender	
		Female	Male
Employee Won	Count Percent	59 (61.5)	317 (73.0)
Employee Lost	Count Percent	37 (38.5)	117 (27.0)

Controlling for applicant gender eliminates any statistically significant adjudicator gender effect. In other words, among female applicants for example, there was no greater likelihood of winning with a male adjudicator than with a female adjudicator. Likewise for male applicants as a subsample. Controlling for adjudicator gender, female applicants were more likely than male applicants to have won with both male adjudicators and female adjudicators, though the relationship is marginally stronger for female adjudicators.

The frequencies for compensation outcomes to successful applicants and applicant and adjudicator gender are set out in Tables Four and Five.

Table Four: Compensation Outcomes by Applicant Gender

		Applicant Gender	
		Female	Male
No Award	Count Percent	11 (8.0)	29 (12.2)
$1 - 5,000	Count Percent	97 (70.3)	125 (52.5)
$5,001 - $10,000	Count Percent	23 (16.7)	60 (25.2)
$10,000 and above	Count Percent	7 (5.1)	24 (10.1)
TOTALS		138	238

Table Five: Compensation Outcomes by Adjudicator Gender

		Applicant Gender	
		Female	Male
No Award	Count Percent	12 (20.3)	28 (8.8)
$1 - 5,000	Count Percent	35 (59.3)	187 (59.0)
$5,001 - $10,000	Count Percent	10 (16.9)	73 (23.0)
$10,000 and above	Count Percent	2 (3.4)	29 (9.1)
TOTALS		59	317

Again, at the level of correlation analysis, compensation awarded under the "humiliation, loss of dignity, and injury to feelings" head showed gender patterns, more so as related to applicant gender (Chi-square $<.01$). Successful male applicants were more likely than successful female applicants to have been awarded no compensation under this head, but those who were awarded compensation were likely to have been awarded more than their female counterparts. The relationship to adjudicator gender was also statistically significant, although less so (Chi-square $<.05$). Female adjudicators were more likely than male adjudicators to have made no compensation award to successful applicants, and less likely than male adjudicators to have awarded compensation awards in the higher ranges, but the differences appear even on the face of the table to be quite marginal.

In any event, in correlational analysis, both gender variables showed relationships to both of our outcome variables. That, however, is just the beginning of the story. In Chi-square tests for correlations, several of the other independent variables showed up as significantly related to the outcome variables as well, suggesting to us the need for a somewhat more sophisticated statistical analysis of causal relationships.

First, applicants in our sample who were dismissed for stated reasons of misconduct were far less likely to have had their grievances sustained by the adjudicator than those dismissed for performance or redundancy reasons, with applicants dismissed for performance reasons most likely to have won in adjudication. On the other hand, reason for dismissal showed no significant relationship to compensation awarded to successful applicants.

Second, applicants who had had their cases adjudicated within the jurisdiction of the Auckland Tribunal were significantly less likely to have won their cases than applicants bringing their cases in the Wellington and Christchurch jurisdictions. Jurisdiction also bore a relationship to compensation, with successful applicants in the Auckland jurisdiction somewhat more likely than those in the Christchurch and, particularly, Wellington jurisdictions to have received a low compensation award.

Third, applicants in our white collar and blue collar categories were about equally successful in winning at adjudication, but they were somewhat less likely to have been successful than those applicants in the managerial and pink collar categories. In terms of compensation to successful applicants, those in the managerial category were far more likely than those in the pink and blue collar categories, and somewhat more likely than those in the white collar category, to have received an award towards the higher end of the usual range.

Only the two representation variables – whether the applicant on the one hand, and the respondent on the other, were represented by lawyers or non-lawyers – showed no significant relationship to the win-lose outcome dimension in correlation tests. Curiously, however, respondent representation showed a statistically significant relationship, albeit a marginal one, to compensation awards to successful applicants. While the pattern was fairly complicated, most obviously respondents employing lawyers at hearing were some-

what more successful in having applicants denied a compensation award despite having had their grievances sustained.

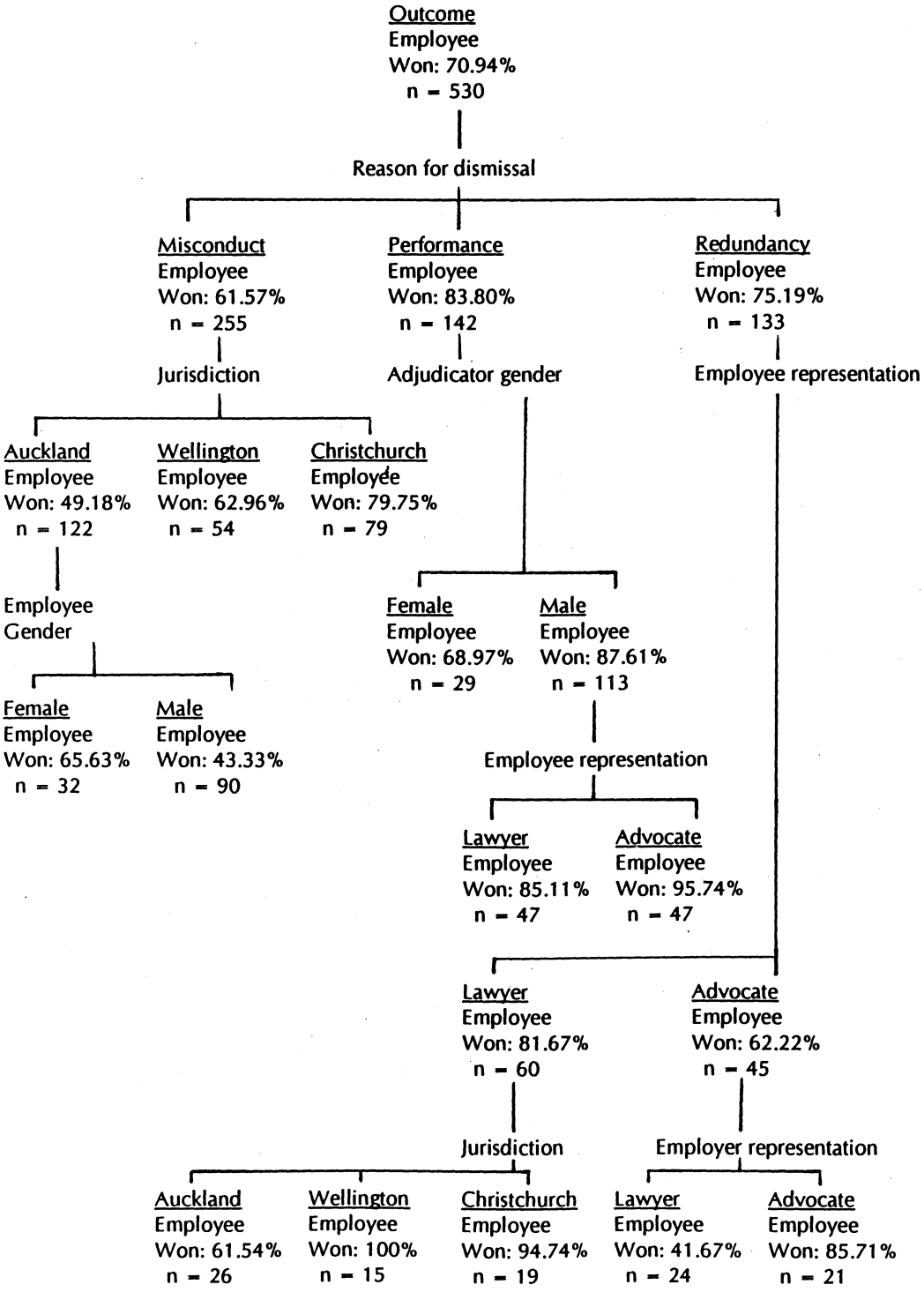
Gender factors in perspective

To sort through the relationships of these several variables to the two outcome measures, we subjected the data to CHAID analysis. CHAID (Chi-squared Automatic Interaction Detector) divides a population into two or more distinct groups based on categories of the "best" predictor of a dependent variable. It then splits each of these groups into smaller subgroups based on other predictor variables. This splitting process continues until no more statistically significant predictors can be found (Magidson, 1992). A CHAID display takes the form of a tree diagram. Figure One is a diagrammatic representation of the variables predicting an employee win in adjudication within our particular sample.

There are a number of interesting predictive relationships suggested by Figure One, although anything more than a passing comment on most of them is beyond the immediate scope of this paper. Clearly, however, at least within our sample of decisions adjudicated in 1993 - 1994 on personal grievance dismissals for cited reasons of misconduct, poor performance or redundancy, the most significant predictor of outcome is the reason for dismissal. The occupational dimension does not feature at all. The gender, jurisdiction and representation variables show up in an interesting mosaic. Amongst decisions involving dismissal for redundancy, the predictors are largely the representation variables, perhaps reflecting what might be seen as the relative complexity of the law surrounding redundancy dismissals. Jurisdiction shows up as a predictor in the cases of employees dismissed for redundancy and where the applicant is represented at hearing by a lawyer, with applicants in this subsample somewhat more likely to have been successful in the Wellington and Christchurch Tribunals. There could, of course, be any number of reasons for this pattern, and it should in any event be noted that the case numbers at that level on the chart are quite small.

Gender factors did show up as predictors in the misconduct and performance subsamples. In the case of misconduct dismissals, jurisdiction is the next best predictor of applicant success, with applicants somewhat more likely to have been successful in the Wellington, and particularly the Christchurch Tribunals than in the Auckland Tribunal. Again, there can be any number of explanations for this pattern, but an indepth analysis is well beyond the scope of this paper. It is possible that in the larger cities less meritorious or marginal cases are more likely to be taken to the Tribunal than would be the case in less urban areas, or that there is a greater inclination to pursue less meritorious cases through to adjudication. If only by virtue of numbers, the quality of representation may be more variable in larger centres. And there are speculatively many other possible explanations for geographical patterns that will have to await future detailed analysis.

Figure One: Variables predicting employee adjudication win

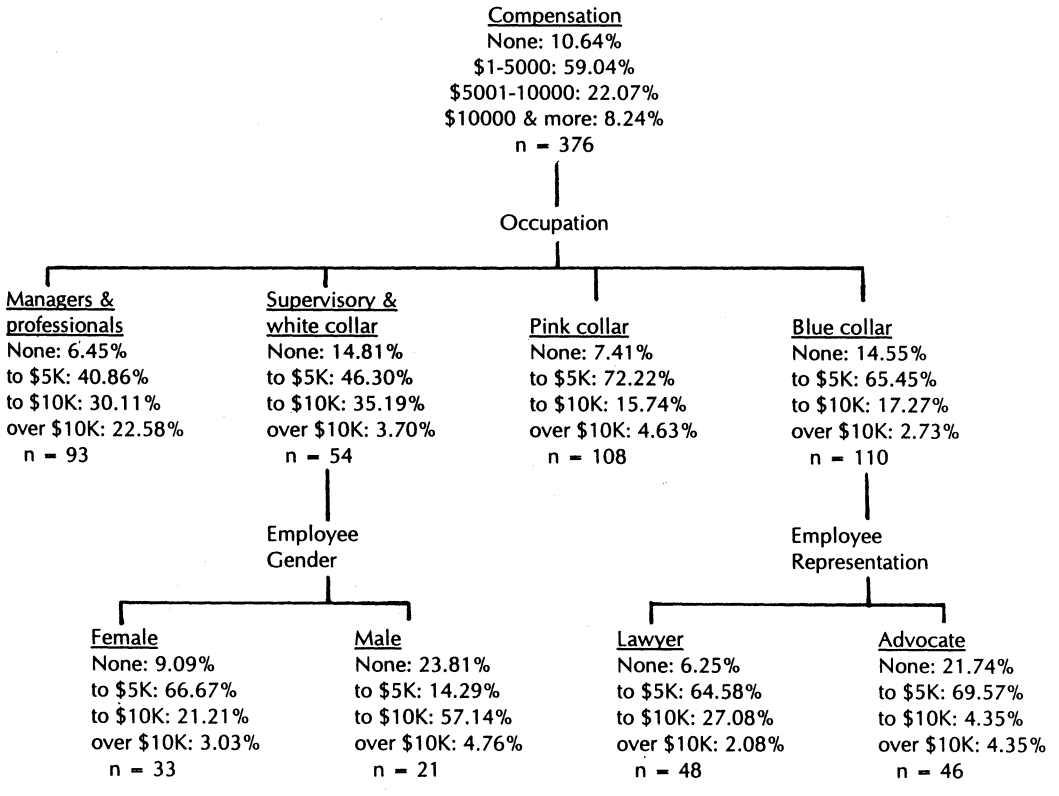


Within the Auckland jurisdiction, amongst applicants dismissed for misconduct, whose cases were decided in adjudication in 1993 - 1994, female applicants were more likely to have won their cases (65.63 percent of 32 applicants) than male applicants (43.33 percent of 90 applicants). Adjudicator gender was not a factor.

Adjudicator gender showed up as a predictor in the subsample of cases involving dismissal for stated reasons of poor performance, with female adjudicators finding for the employee in 68.97 percent of 29 cases and male adjudicators finding for the employee in 87.61 percent of 113 performance cases. Applicant gender was not a factor. The nature of the employee's representation featured as a further predictor in the case of male adjudicators, but not in the case of female adjudicators.

Turning now to the second outcome variable, Figure Two provides a diagrammatic representation of the variables predicting s.40(1)(c)(i) compensation outcomes to successful applicants.

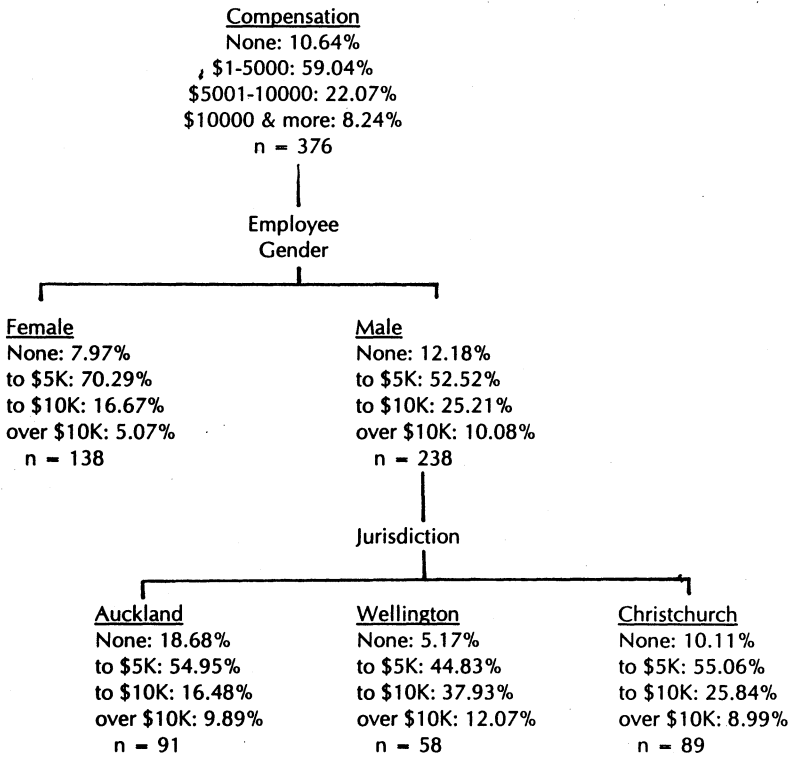
Figure Two: Variables predicting s. 40 (1) (c) (i) compensation



Here, while the patterns are not uncomplicated, occupation is clearly the dominant predictor of compensation outcomes for successful applicants in our sample. It is difficult to read much into the "no award" figures. But most obviously, applicants in the pink collar and blue collar categories were more likely than those in the two white collar categories to be awarded compensation of $5,000 or less, and less likely to be awarded compensation over $10,000. Those in the managerial and professional category were, by a considerable margin, more likely than others to be awarded over $10,000. Employee gender is a predictor within the supervisory and white collar category, with male applicants more likely to have received higher awards.

Removing the occupational variable from the equation produces the diagrammatic result in Figure Three.

Figure Three: Variables (except occupation) predicting compensation



While Figure Three results from an artificial statistical manipulation in omitting the all important occupation variable, it does give some indication of the gender compen-sation pattern that appears to be generated by occupational gender patterns, coupled perhaps with compensation tied in part to applicant pay rates. With occupation taken out of the equation, only gender shows up as a substantial predictor of compensation, with some much smaller variations by jurisdictions. It is noteworthy that the overall gender pattern differences in Figure Three are far less sizeable than those within the supervisory and white collar category in Figure Two, perhaps confirming our suspicion that the overt gender factor in Figure Two is really masking an extended occupational effect within that category.

Discussion and conclusions

We began by briefly reviewing theories that have informed the overseas research into gender influences in employment arbitration. The chivalry/paternalism thesis suggested that male arbitrators or adjudicators are more lenient to female grievants than to male grievants. The evil woman thesis suggested something of the opposite: that male adjudicators are harsher on female grievants than on male grievants, though the reach of this theory is really limited to "sentencing" or the punishment imposed on grievants found to be at fault. The queen bee thesis suggests that female adjudicators discriminate against women grievants.

We reviewed the overseas research and found three general themes emerging from the mixed results of that body of work. First, the centrepiece of research on gender and employment arbitration, the Bemmels studies in the late 1980s and early 1990s, generally found evidence to support the chivalry/paternalism thesis. Across his several studies, Bemmels often found evidence of more favourable treatment of women grievants than of men grievants by male arbitrators. Second, while some other studies registered findings similar to Bemmels', there were still others that showed no patterns of arbitration outcomes by grievant gender at all. The third theme that emerged from our review of the previous research was the almost unanimous finding of no relationship between arbitrator or adjudicator gender and arbitration outcomes.

We also reviewed the New Zealand study by Morris which reported, relevantly to this report, no evidence of gender bias in win - lose decisions in the Employment Tribunal, but evidence of occupationally related gender patterns, disadvantaging successful female grievants, in compensation awards for humiliation, loss of dignity and injury to feelings.

From the body of theory and research we proposed three hypotheses for testing: that male adjudicators treat female grievants more favourably than male grievants; that female adjudicators treat male grievants more favourably than female grievants; and that female adjudicators treat grievants more favourably than do male adjudicators.

What we found was, first, that for the subsample of applicants dismissed for misconduct and whose cases were decided in the Auckland Tribunal, female grievants were more likely than male grievants to have won their cases, regardless of the gender of the adjudicator.

Second, for the subsample of applicants dismissed for performance reasons, male adjudicators were more likely to have found for the applicants than female adjudicators, regardless of the gender of the applicant. And third, there were clear gender patterns to compensation awards, though the evidence supports a finding that those patterns derive from occupational gender patterns.

And so, we found no support for any of the hypotheses as such, and there was some evidence in the direction contrary to Hypothesis Three. It should be said that a real test of the evil woman thesis would probably require examining "sentencing decisions," most obviously some aspects of remedies such as reduction of remedies for contributory conduct. There was some support of a sort for Hypothesis One in the Auckland misconduct subsample, in that female applicants did better than male applicants with male adjudicators. But they also did better with female adjudicators.

In terms of our first outcome, the win - lose measure, there was then some very limited evidence of gender factors perhaps at work, but not in the ways predicted by theory. And it is noted that the gender factors that did show up were buried at levels two and three of our CHAID regression diagram. Reason for dismissal was clearly the first and best predictor of whether the applicant was likely to win or lose.

The most apparent gender pattern that we can point to in this report is the gender pattern to compensation under the humiliation, loss of dignity and injury to feelings head, though it is a derived pattern. Our results endorse the findings of Morris (1996) in this respect. The gender pattern showed up most starkly in the CHAID analysis once the occupation variable was removed, but was invisible except in the supervisory and white collar occupational group with the occupation variable included. Given the dichotomous nature of the supervisory and white collar category, we feel comfortable in concluding that the apparent gender factor within that occupational category in fact masked a further occupational influence.

It is clear, then, that the apparent gender pattern to compensation awards that shows up in our sample derives from occupational gender patterns and the apparent reference to applicant income as one influence on compensation award levels. Nonetheless, though it be effect rather than cause, the gender pattern to compensation, if shown to be sustained over time, might give reason for review.

It is appropriate to reiterate several caveats before closing. To begin with, the sample is a relatively small sample from the population of Employment Tribunal adjudications since the inception of the Tribunal in 1991, and an even smaller sample of total case applications and dispositions since inception. Accordingly, the results cannot necessarily be extrapolated beyond the sample. Several years have passed since the decisions analysed for this report were issued. Aspects of the law that guide the Tribunal have been refined or more closely defined over that time, including through a number of important decisions of the Employment Court addressing factors that are appropriate and inappropriate for consideration in the setting of compensation levels.

And, of course, our findings rely on a number of untested assumptions, including most critically the assumption that the 530 cases in the sample are of equal merit, without any differentiation across gender, reason for dismissal, occupation, geographic jurisdiction or representation type. That assumption, in turn, rests on further assumptions, for example an equal propensity to mediate rather than adjudicate across applicant gender, reason for dismissal, representation type and so forth. In examining patterns of compensation awards, we have not factored in other remedies awarded to the successful applicants. So, again, it is appropriate to be cautious in drawing any implications beyond the limits of the subsample reported here.

Having said that, it can certainly also be said that there are topics for future examination suggested by this material. The study of the gender factor itself can be brought up to date and taken to more sophisticated levels. Its side trails might also be interesting to explore. In our Figure One, for example, the nature of the grievant's representation showed up as of third-level significance in certain circumstances where the adjudicator was male, but not where the adjudicator was female. That, too, is an aspect of the gender factor, not necessarily a critical one, but nonetheless one of both academic and practical curiosity. And, of course, the factors beyond gender that seemed in some circumstances to be associated with adjudication outcomes in our sample will be examined more closely in future work.

References

Bemmels, Brian (1988a) The Effect of Grievants' Gender on Arbitrators' Decisions *Industrial and Labor Relations Review* **41(2):** 251-262.

Bemmels, Brian (1988b) Genders Effects in Discharge Arbitration *Industrial and Labor Relations Review* **42(1):** 63-76.

Bemmels, Brian (1988c) Gender Effects in Discipline Arbitration: Evidence from British Columbia *Academy of Management Journal* **31(1):** 699-706.

Bemmels, Brian (1990) Arbitrator Characteristics and Arbitrator Decisions *Journal of Labor Research* **11(2):** 181-192.

Bemmels, Brian (1991) Gender Effects in Grievance Arbitration *Industrial Relations* **30(1):** 150-162.

Bigoness, William J. and Philip B. DuBose (1985) Effects of Gender on Arbitrators' Decisions *Academy of Management Journal* **28(2):** 485-491.

Caudill, Steven B. and Sharon L. Oswald (1992) An Alternative to Bemmels' Method of Investigating Biases in Arbitration. Brian Bemmels Reply. *Industrial and Labor Relations Review* **45(4):** 800-805.

Caudill, Steven B. and Sharon L. Oswald (1993) A Sequential Selectivity Model of the Decisions of Arbitrators *Managerial and Decision Economics* **14(3):** 261-267.

Crow, Stephen M. and James W. Logan (1994) Arbitrators' Characteristics and Decision-Making Records, Gender of Arbitrators and Grievants, and the Presence of Legal Counsel as Predictors of Arbitral Outcomes *Employee Responsibilities and Rights Journal* **7(2):** 169-185.

Davis, Wendy (1994) *A Feminist Perspective on Sexual Harassment in Employment Law in New Zealand* Monograph. Wellington: New Zealand Institute of Industrial Relations Research.

Magidson, Jay (1992) *SPSS CHAID Version 5*. Chicago: SPSS Inc.

Morris, Caroline (1996) An Investigation into Gender Bias in the Employment Institutions *New Zealand Journal of Industrial Relations* **21(1):** 67-90.

Moulds, Elizabeth F. (1980) Chivalry and Paternalism: Disparities of Treatment in the Criminal Justice System. In Susan K. Datesman and Frank R. Scarpitti (eds) *Women, Crime and Justice* New York: Oxford University Press; pp. 277-299.

Oswald, Sharon L. and Joseph G. VanMatre (1990) Arbitration and the Queen Bee Syndrome *Review of Business and Economic Research* **26(1):** 38-45.

Ponak, Alan (1987) Discharge Arbitration and Reinstatement in the Province of Alberta *The Arbitration Journal* **42(2):** 39-46.

Rodgers, Robert C. and I. B. Helburn (1984) The Arbitrariness of Arbitrators' Decisions *Proceedings of the 37th Annual Meeting of the IRRA* (Madison, WI: IRRA); pp. 234-241.

Rosenberg, Sheldon (1979) Sex Discrimination in the Labor Arbitration Process *Labor Law Journal* **30:** 102-117.

Scott, Clyde and Elizabeth Shadoan (1989) The Effect of Gender on Arbitration Decisions *Journal of Labor Research* **10(4):** 429-436

Steen, Jack E., Pamela L. Perrewe and Wayne A. Hockwarter (1994) A Re-examination of Gender Bias in Arbitration Decisions *Labor Law Journal* **45:** 298-305.

Thornton, Robert J. and Perry A. Zirkel (1990) The Consistency and Predictability of Grievance Arbitration Awards *Industrial and Labor Relations Review* **43(2):** 294-307.

Zirkel, Perry A. (1983) A Profile of Grievance Arbitration Cases *The Arbitration Journal* **37:** 35-38.