FEDERAL JOB CLASSIFICATION
Comparison of Job Content With Grades Assigned in Selected Occupations

November 1995

United States General Accounting Office
Report to Congressional Requesters

GAO/GGD-96-20
Over the years, many studies have suggested that women and minorities are paid less than men and nonminorities who work in comparable positions. These observations have raised questions about whether the federal government’s classification systems result in lower grades being assigned to positions in occupations having large numbers of female or minority incumbents than to other comparable occupations. This report responds to your request that we determine whether the relationship between job content and grades assigned using the Factor Evaluation System (FES) varied on the basis of the proportions of women and minorities in occupations.¹

Background

Position classification systems are formal methods for determining the relative worth of positions in an organization. The Office of Personnel Management (OPM) has responsibility and authority for federal position classification, except for certain positions in agencies exempted by law. OPM develops and issues classification standards and policies for the federal personnel system. Federal agencies then use these standards and policies to assign grades to positions.

The General Schedule (GS) is a 15-grade pay system that covers 442 white-collar occupations and approximately 1.5 million full-time

¹FES, which covers almost one-third of the federal full-time permanent white-collar workforce, is a point-factor system for determining grades and is considered to be more orderly or rigorous than other federal classification systems.
permanent employees in the federal white-collar workforce. Agencies classify most of these positions using either narrative or FES point-factor classification standards. The FES primary standard serves as the framework for individual classification standards written for each occupation. When we began our study, FES standards were in effect for 77 occupations covering approximately 441,000 full-time permanent nonsupervisory employees.

Under FES, positions are assigned grades on the basis of the duties, responsibilities, and the qualifications required in terms of the following nine factors:

- **knowledge required by the position** includes the skills needed to apply that knowledge;
- **supervisory controls** entail the control exercised by the incumbent’s supervisor (not the incumbent’s span of control over subordinates);
- **guidelines** provide reference data or impose certain constraints on the incumbent’s use of knowledge; include desk manuals, established procedures and policies, traditional practices, and materials such as dictionaries, etc.; and vary by specificity, applicability, and availability;
- **complexity** consists of the nature, variety, and intricacy of tasks and the difficulty and originality involved in performing the work;
- **scope and effect** encompass the relationship between the breadth and depth of the work and the effect of work products and services within and outside the organization;
- **personal contacts** refer to contacts with persons not in the supervisory chain;
- **purpose of contacts** ranges from factual exchanges of information to situations involving significant or controversial issues and differing viewpoints, goals, or objectives;
- **physical demands** include physical characteristics such as agility, dexterity, and physical exertion—such as climbing, lifting, pushing, etc.; and
- **work environment** pertains to the risks and discomforts in the employee’s physical surroundings or work assigned.

As shown in table 1, each factor is broken down into graduated levels. Factors are composed of from three to nine levels with most having four

---

2The point-factor system involves assigning a point value or weight to each compensable factor and totaling the points assigned to obtain a job worth score that measures the relative importance of each position to an organization. Narrative standards use fewer factors to describe the important characteristics of the work, and the GS grade is determined through nonquantitative analysis.
to six levels. One factor—knowledge required by the position—has the largest range of points, from 50 for level 1 to 1,850 for level 9.

Table 1: FES Factors and Points by Factor Levels

<table>
<thead>
<tr>
<th>Factor</th>
<th>Level²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Knowledge required by the position</td>
<td>50</td>
</tr>
<tr>
<td>Supervisory controls</td>
<td>25</td>
</tr>
<tr>
<td>Guidelines</td>
<td>25</td>
</tr>
<tr>
<td>Complexity</td>
<td>25</td>
</tr>
<tr>
<td>Scope and effect</td>
<td>25</td>
</tr>
<tr>
<td>Personal contacts</td>
<td>10</td>
</tr>
<tr>
<td>Purpose of contacts</td>
<td>20</td>
</tr>
<tr>
<td>Physical demands</td>
<td>5</td>
</tr>
<tr>
<td>Work environment</td>
<td>5</td>
</tr>
</tbody>
</table>

²The blank spaces in columns indicate that no further levels exist for the corresponding factor.

Source: FES Primary Standard.

To determine a position’s GS grade, the agency typically compares either the position description or information gathered through a “desk audit” with the nine FES factors described in the classification standard.³ After all nine factors are evaluated, the points for all factors are totaled, and the total for each position is converted to a GS grade by using a conversion table (see table 2).

³Desk audits typically consist of a trained classifier’s face-to-face interviews with the incumbent and supervisor and observations of the work operations, sometimes including an examination of work products.
Table 2: FES GS Grade Conversion

<table>
<thead>
<tr>
<th>GS grade</th>
<th>Point range</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS-1</td>
<td>190 - 250</td>
</tr>
<tr>
<td>GS-2</td>
<td>255 - 450</td>
</tr>
<tr>
<td>GS-3</td>
<td>455 - 650</td>
</tr>
<tr>
<td>GS-4</td>
<td>655 - 850</td>
</tr>
<tr>
<td>GS-5</td>
<td>855 - 1,100</td>
</tr>
<tr>
<td>GS-6</td>
<td>1,105 - 1,350</td>
</tr>
<tr>
<td>GS-7</td>
<td>1,355 - 1,600</td>
</tr>
<tr>
<td>GS-8</td>
<td>1,605 - 1,850</td>
</tr>
<tr>
<td>GS-9</td>
<td>1,855 - 2,100</td>
</tr>
<tr>
<td>GS-10</td>
<td>2,105 - 2,350</td>
</tr>
<tr>
<td>GS-11</td>
<td>2,355 - 2,750</td>
</tr>
<tr>
<td>GS-12</td>
<td>2,755 - 3,150</td>
</tr>
<tr>
<td>GS-13</td>
<td>3,155 - 3,600</td>
</tr>
<tr>
<td>GS-14</td>
<td>3,605 - 4,050</td>
</tr>
<tr>
<td>GS-15</td>
<td>4,055 - up</td>
</tr>
</tbody>
</table>

Source: FES Primary Standard.

Approach

A contractor, with our supervision, developed a job content questionnaire on the basis of the FES primary standard. The contractor distributed it to a stratified random sample of 2,060 pairs of incumbents and their supervisors and received responses from 1,639 incumbent/supervisor pairs, which represents an overall response rate of about 80 percent. Because individual federal positions are classified through labor-intensive desk audits, it was not practical for us to study the majority of FES occupations using traditional classification methods.

The contractor did, however, do desk audits of 78 judgmentally selected positions and compared the audit results with the related questionnaire responses. These comparisons indicated a fairly high correlation across occupations. The validity coefficient between the GS grades resulting from the desk audits and those from the questionnaires was .80 when the incumbent and supervisor questionnaire responses were averaged.4 We considered this correlation to be sufficiently high to validate the use of the

---

4A validity coefficient measures the consistency between two data sets.
job content questionnaire results for comparing groups of occupations. However, we cannot attest to the questionnaire’s validity when used across GS grades within occupations, for specific occupations, or for individual positions. Further, our study was not designed to determine whether, and if so why, the job content reflected by the questionnaires differed from that contained in the official job descriptions, which are the bases for the actual GS grades.

We used an OPM database to identify a representative segment of incumbents and their supervisors in occupations with high, medium, or low representations of women or minorities. We defined occupations with high, medium, and low female representation as those in which women represented 70 percent or more, 31 to 69 percent, and 30 percent or less of incumbents, respectively. We considered occupations in which minorities represented more than 41 percent, 23 to 41 percent, or less than 23 percent of incumbents as those with high, medium, and low minority representation, respectively. Examples of occupations with high female representation included secretary, dental hygienist, medical clerk, dental assistant, and occupational therapist. Examples of occupations with high minority representation included border patrol agent and computer and equal employment opportunity specialists.

We selected our sample of incumbents and their supervisors from a total of 58 occupations, which collectively represented about 90 percent of the full-time permanent nonsupervisory employees covered by FES when we designed our study. Because preliminary analyses indicated that more variation in undergrading and overgrading existed among occupations with similar gender and minority representation than between groups of occupations with different gender and minority representations, we did our analyses on the 37 occupations for which we had received completed questionnaires from at least 10 or more incumbent/supervisor pairs, for a total of 1,358 pairs or positions. These 37 occupations represented almost one-quarter of the federal white-collar workforce, or about 79 percent of the employees covered by FES, and the results of our study are generalizable only to this population.6

5Psychological testing standards traditionally required validity coefficients of at least .70 for all tests, and on that basis, we conclude that our result is sufficiently valid. However, psychologists have acknowledged that coefficients as low as .30 can be of practical value. This is the only classification study we are aware of in which desk audits were used to demonstrate that the questionnaire results were similar to those which would be attained in an actual position classification.

6Because we could not obtain reliable estimates of the odds of overgrading and undergrading for occupations with small numbers of respondents, we deleted those 21 occupations with fewer than 10 incumbent/supervisor pairs.
On the basis of the average total factor level points derived from the questionnaires completed by the incumbent and the supervisor, we determined the GS grade for each position using the FES conversion table (see table 2). We compared the questionnaire grade with the incumbent’s actual GS grade and considered the position to be aptly, or appropriately, graded if the questionnaire grade and the actual grade were the same. Otherwise, we considered positions to be overgraded if the actual grade was higher than the questionnaire grade and undergraded if the actual grade was lower than the questionnaire grade.

Our study was not designed to permit us to

- approximate the number of positions that appeared to be overgraded, undergraded, or aptly graded for the portion of the federal workforce to which the results of our study are generalizable;
- identify the causes of any overgrading or undergrading resulting from either (1) our use of the primary rather than the occupation-specific classification standards, (2) the agencies’ application of classification standards to individual positions, or (3) management decisions regarding the work incumbents were actually assigned versus their job descriptions;
- determine whether any difference on the basis of gender or minority status was inherent in the design of FES, as a product either of the factors that constitute FES or the allocation of weight or the point range assigned to each factor; or
- calculate what pay adjustments, if any, should be made.

Appendix I contains a more detailed discussion of the job content questionnaire development and validation, sample selection, and response rate calculation.

To determine the effects of female or minority representation on relative overgrading and undergrading, we used odds and odds ratios. We calculated the odds of occupations being undergraded rather than aptly graded by dividing the number of positions undergraded by the number aptly graded for groups of occupations. To determine how much more likely one group of occupations was to be undergraded than another group, we divided the odds of being undergraded for one group of occupations by the odds for the other group to form an odds ratio. We used the same procedures for overgrading.

We used loglinear analysis to determine how the odds of occupations being overgraded or undergraded versus aptly graded varied (1) across the
range of GS grades and (2) when the female or minority representation of the occupation was high, medium, or low. The strength of this particular statistical approach is that multiple variables can be analyzed simultaneously. Appendix II provides more detailed information on the calculation of odds and odds ratios and loglinear models tested and the results obtained.

To gain historical perspective, we reviewed previous studies of federal classification issues. We also conferred with federal classification experts to obtain additional insights about possible explanations for specific findings.

This study should not be referred to as a “pay equity” study because we examined only the relationship between job content and GS grades assigned through the use of FES and whether that relationship varied with the proportion of women or minorities in occupations.

We obtained comments from OPM that are discussed on pages 11 through 12 and presented in appendix III. We did our study from January 1990 to September 1995 in accordance with generally accepted government auditing standards.

Results in Brief

For the nonsupervisory positions in the 37 occupations we studied, our analyses suggest that the difference between the actual GS grades and those we determined using a job content questionnaire was directly related to the female and minority representation of the occupations we examined. If the actual grade was higher than the questionnaire-derived grade, we considered the position to be overgraded; if the actual grade was lower than the questionnaire grade, we considered the position to be undergraded.

The likelihood of a position being overgraded, rather than aptly graded, increased as the incumbents’ GS grades increased. However, the incumbents’ grades had virtually no effect on the likelihood that a position was undergraded versus aptly graded. After statistically eliminating this effect of the incumbents’ GS grades, we found that the occupations we

---

7We developed our job content questionnaire on the basis of the FES primary standard that serves as the framework for individual classification standards written for each occupation. The actual GS grades were assigned using these occupation-specific standards. Further, the questionnaire was designed and validated to achieve our review objectives relative to comparing groups of occupations. It was not designed as and is not a valid substitute for traditional classification methods. Therefore, the questionnaire should not be used to draw definitive conclusions regarding overgrading or undergrading of individual positions.
studied with high female representation were more likely to be undergraded rather than aptly graded compared with the occupations having medium or low female representation. As mentioned above, we defined occupations with high female representation as those in which 70 percent or more of the incumbents were women. This definition was used on the basis of the current literature when we designed our study.

The literature contained no consistent basis on which to define minority representation. On the basis of our past work, we initially adopted a standard of more than 48 percent to define high minority representation. However, we found it necessary to adjust our definition for high minority representation to occupations in which more than 41 percent of the incumbents were minorities in order to include at least two occupations with each possible mix of gender and minority representation (e.g., low female, high minority representation).

We would have preferred that our original definition had provided a sufficient mix of occupations in our sample. And it is important to note that about half of the incumbents in the total population of the occupations defined by our study as having a high minority representation were in fact nonminorities. Within these parameters, after eliminating the effects of the incumbents’ GS grades, we found that occupations with high minority representation were more likely to be overgraded rather than aptly graded in comparison with the occupations having medium or low minority representation.

The National Performance Review has contended that the current federal classification systems have too many occupations and grades and has recommended that a more flexible “broad-banded” system be adopted. OPM is currently, within the existing statutory framework, planning to revise the classification standards and increase classification oversight. Our study suggests any new system should be closely monitored to ensure that unintended disparities are identified and addressed.

---

8When we designed our study, minorities comprised approximately 32 percent of the workforce covered by FES. We initially defined occupations with high minority representation as those in which more than 48 percent (150 percent of the 32 percent minority workforce representation) of the incumbents were minorities.
Effects of Female and Minority Representation on Relative Overgrading and Undergrading

To evaluate whether gender or minority representation had an effect on the variation between the actual grades and the questionnaire grades, we statistically eliminated the effect of the incumbents’ GS grades. Analysis of the remaining variation showed that occupations with high female representation were 1.77 times more likely to be undergraded rather than aptly graded compared with occupations having a low or medium female representation. Occupations with high minority representation were 2.18 times more likely to be overgraded rather than aptly graded compared with occupations having a low or medium minority representation.

Classification experts with whom we consulted about our results and the available literature offered a few occupation-specific hypotheses about possible causes. For example, key occupations with high minority representation that appeared to be overgraded included (1) border patrol agents, (2) equal employment opportunity and compliance specialists, and (3) computer specialists. Previous studies of federal classification issues maintained that FES was ineffective for specialists such as law enforcement related occupations because physical demands and work environment are not highly valued FES factors but are considered significant in these occupations.

Although empirical data are lacking, the classification experts we consulted suggested that when equal employment opportunity occupations were established in the 1970s, they involved a heavy workload of cases and, even though not recognized by FES, the GS grades of these occupations may have been increased on that basis. Furthermore, private sector wages may have resulted in overgrading positions in computer-related occupations. Explanations are somewhat less evident regarding occupations with high female representation, which appear more likely to be undergraded.

---

9The likelihood or odds of a position being overgraded increased as the GS grade increased; that is, as the grade increased by one GS grade, the odds that a position was overgraded versus aptly graded increased by a factor of 1.64. For example, the odds of positions being overgraded in occupations with an average grade of GS-9 were approximately 5 to 1 (i.e., five positions were overgraded for every one that was aptly graded), while the odds of positions being overgraded in occupations with an average grade of GS-10 were about 8.2 to 1. The factor of 1.64 (8.2 / 5 = 1.64) indicates that positions in the higher graded occupations were 1.64 times more likely than those in the lower graded occupations to be overgraded. At the 95 percent level of confidence, as the grade increased by one GS grade, the odds that a position was overgraded versus aptly graded increased by a factor ranging from 1.53 to 1.77.

The GS grade had no significant effect on the odds that a position was undergraded versus aptly graded.

10At the 95 percent level of confidence, occupations with high female representation were from 1.33 to 2.35 times more likely to be undergraded than occupations with medium or low female representation, and occupations with high minority representation were from 1.62 to 2.94 times more likely to be overgraded than those with medium or low representation.
Current Efforts to Revamp Classification Standards and Systems

Critics of the federal classification system, including reports from the National Academy of Public Administration and the National Performance Review, have argued that the classification system is overly complex, with too many occupations and GS grades, and that a less rigid system would be more effective. They have recommended fewer occupations and more flexible broad-banded grade structures.

OPM officials told us recently that, within the existing statutory framework, OPM plans to revise the classification standards and increase its oversight of various processes including classification. A current proposal for rewriting classification standards would reduce the inventory of 442 white-collar classification standards to about 74. Rather than individual occupations, the new standards would focus on the 22 “job families” of related occupations with separate standards, as applicable, for professional, administrative, technical, and clerical positions.11

OPM has also established a new oversight office, which, among other things, is planning various governmentwide policy studies. OPM has tentatively allocated about 145 staff years to this effort; most of these resources are located in field offices rather than at headquarters. One of the highest priorities will be a governmentwide classification study, with particular emphasis on determining the accuracy of “border grades”—those grades most likely to be placed at the lower and upper limits of any newly created grade bands. A team is examining options for doing this study, and work on the study is scheduled to begin early in fiscal year 1996.

Conclusions

Although FES is considered a more orderly or rigorous method than other federal classification systems, our study identified differences in the grading of positions in occupations with high representations of women or minorities. The National Performance Review and other studies suggest that the current classification systems should be abandoned in favor of more flexible, broad-banded systems. The results of our study indicate that it is important that policymakers closely monitor any new systems to ensure that (1) unintended disparities are identified so that they can be corrected and (2) the national policy underlying the current classification system—that jobs be classified so that pay is equal for substantially equal work—is being satisfactorily achieved. Since OPM is in the process of

---

11“Job families” consist of related occupations that are grouped together such as the General Administrative, Clerical, and Office Service Group; the Accounting and Budget Group; the Physical Science Group; etc.
substantially revising the classification and oversight systems, we are making no recommendations in this report.

Agency Comments and Our Evaluation

OPM provided written comments on a draft of this report. (See app. III.) OPM took issue with our methodology, saying that it was insufficient to support our findings. OPM also discussed its plans to further explore job classification accuracy issues.

OPM took exception with our methodology in two respects. First, it believes we inappropriately used the primary classification standard rather than occupation-specific standards as the basis for our job content questionnaire. OPM said that the primary standard was designed to be used as an overall outline wherein more specific standards would be developed, not as a basis for evaluating individual positions. Although we used the primary standard rather than the occupation-specific standards in the development of our methodology, we examined the specific standard for several occupations in our sample to see if we could identify any ways in which the use of the occupation-specific standard might have led to a different result. We did not identify any such effect. We also asked OPM's classification experts to identify occupations in our sample for which the specific standards were, in their view, sufficiently different from the primary standard that our results would have been affected in identifiable ways. They did not identify any such occupations. Thus we continue to believe that our use of the primary standard was appropriate and that our methodology produced useful results.

OPM's second exception with our methodology was our use of a job content questionnaire rather than traditional desk audits to assign grades to positions. OPM said that our use of a questionnaire resulted in employees and supervisors, unfamiliar with FES ground rules, being asked to select generic phrases that were not in context and that this in turn resulted in the grades we assigned being less credible than those derived by federal agencies. We acknowledge in our text that actual GS grades are assigned on the basis of occupation-specific standards and that the desk audit is a typical way to assign a grade to a position. As noted in the text, desk audits are labor-intensive, and it was not practical for us to study the majority of FES occupations using traditional classification methods. Because of this, we took care to validate our results. First, we had a contractor do desk audits on positions in a number of occupations in our sample. Next, we compared the results of those desk audits with the questionnaire results for those positions. This comparison showed a fairly
A high correlation between the GS grades resulting from desk audits and those from the questionnaire. Thus, we believe that our methodology is appropriate to identify patterns of overgrading and undergrading among groups of occupations with different representations of women or minorities.

OPM also questioned our study results by comparing them with other studies. More specifically, OPM said that other agencies’ studies and OPM’s classification appeals data indicated lower levels of misclassification than our study. We are unaware of any recent studies or appeals data in which a direct comparison with our study could be meaningful. Although OPM’s most recent report on the overall federal white-collar position classification accuracy indicated a lower level of misclassification than our study, it was published in 1983. We acknowledge that classification appeals also indicate a lower level of misclassification than our study. However, classification experts with whom we consulted said that appeals data are unlikely to represent those federal employees whose positions may be overgraded. As indicated by the Merit System Protection Board (MSPB), almost no one files a classification appeal.12

Finally, OPM said that it shares with us the need to ensure that the federal government’s classification systems and their applications are fair and unbiased. OPM said that to this end, its newly designed oversight program will have a major focus on ensuring that current and new classification systems advance the merit principles of equal pay and the efficient and effective use of the federal workforce. OPM said that it expects to decide on the classification review design by the end of fiscal year 1995 and begin work on the review in early fiscal year 1996.

We are sending copies of this report to interested Members of Congress and congressional committees that have responsibilities for public sector employment issues, the Director of the Office of Personnel Management, and other interested parties. Copies will also be made available to others upon request.

12According to MSPB only 240 of the government's 2.1 million civilian nonpostal employees filed an appeal in 1988, or about one one-hundredth of 1 percent.
Major contributors to this report are listed in appendix IV. If you have any questions about the report, please call me at (202) 512-7824.

Nancy Kingsbury
Director
Planning and Reporting
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Letter</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Appendix I</strong></td>
<td></td>
</tr>
<tr>
<td>Questionnaire Development and GS Grade Estimation</td>
<td>18</td>
</tr>
<tr>
<td>Sample Selection</td>
<td>19</td>
</tr>
<tr>
<td>Response Rate Calculation</td>
<td>24</td>
</tr>
<tr>
<td>Occupations Included in the Study Analyses</td>
<td>33</td>
</tr>
<tr>
<td>Questionnaire Validation</td>
<td>34</td>
</tr>
<tr>
<td><strong>Appendix II</strong></td>
<td></td>
</tr>
<tr>
<td>Loglinear Methodology and Analysis Results</td>
<td>39</td>
</tr>
<tr>
<td>Data Analysis Approach</td>
<td>39</td>
</tr>
<tr>
<td>Occupations Included in Our Study</td>
<td>39</td>
</tr>
<tr>
<td>Relationship Between Overgrading/Undergrading and Female or Minority Representation Without Controlling for GS Grade</td>
<td>45</td>
</tr>
<tr>
<td>Relationship Between Overgrading/Undergrading Versus GS Grade and Female and Minority Representation</td>
<td>50</td>
</tr>
<tr>
<td><strong>Appendix III</strong></td>
<td></td>
</tr>
<tr>
<td>Comments From the Office of Personnel Management</td>
<td>63</td>
</tr>
<tr>
<td><strong>Appendix IV</strong></td>
<td></td>
</tr>
<tr>
<td>Major Contributors to This Report</td>
<td>68</td>
</tr>
<tr>
<td><strong>Bibliography</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Related GAO Products</strong></td>
<td>144</td>
</tr>
<tr>
<td>Tables</td>
<td></td>
</tr>
<tr>
<td>Table 1: FES Factors and Points by Factor Levels</td>
<td>3</td>
</tr>
<tr>
<td>Table 2: FES GS Grade Conversion Table</td>
<td>4</td>
</tr>
<tr>
<td>Table I.1: FES Factors and Points by Factor Levels</td>
<td>20</td>
</tr>
<tr>
<td>Table I.2: Conversion of FES Points to GS Grade</td>
<td>23</td>
</tr>
</tbody>
</table>
Table I.3: Example of GS Grade Estimation Based on Incumbent and Supervisor Questionnaire Responses 24
Table I.4: Distribution of the 77 FES Occupations 25
Table I.5: Female and Minority Representation, PATCO Category, and Number of Incumbents for the 77 Occupations Covered by FES 27
Table I.6: Distribution of the 58 Occupations Included in the Sample 30
Table I.7: Distribution of the 2,233 Incumbents by Female and Minority Representation and Strata 32
Table I.8: Disposition of Incumbent/Supervisor Pairs in the Sample 33
Table I.9: Female and Minority Representation of the 37 Occupations Included in the Study 35
Table I.10: Distribution of Desk Audits Completed According to Female and Minority Representation for the 37 Occupations Included in the Study 36
Table I.11: Validity Coefficients for Grades Estimated on the Basis of the 78 Desk Audits Versus the Questionnaire Responses for Incumbent/Supervisor Pairs, Supervisors Only, and Incumbents Only 38
Table II.1: Female and Minority Representation, Average GS Grades, Observed Sample Frequencies, and Odds of Being Overgraded and Undergraded Versus Aply Graded for 37 Occupations Without Controlling for GS Grade 42
Table II.2: Relationship Between the Observed Sample Frequencies and Female Representation Without Controlling for GS Grade 46
Table II.3: Hierarchical Models Tested to Examine Relationships Between Observed Sample Frequencies and Female and Minority Representation 47
Table II.4: Relationship Between the Expected Frequencies and Female Representation Without Controlling for GS Grade 48
Table II.5: Relationship Between the Observed Sample Frequencies and Minority Representation Without Controlling for GS Grade 49
Table II.6: Relationship Between the Expected Frequencies and Minority Representation Without Controlling for GS Grade 50
Table II.7: Loglinear Models Tested to Examine the Relationship Between the Frequencies, the GS Grade, and Female and Minority Representation 54
Table II.8: Female and Minority Representation, Average GS Grades, Expected Frequencies, and Odds of Overgraded and Undergraded Versus Aptly Graded for 37 Occupations
Table II.9: Coefficients and Odds Ratios Describing the Effects of the Average GS Grade and High Female and Minority Representation on Overgrading and Undergrading

Figures
Figure I.1: Excerpt from the FES Primary Standard
Figure I.2: Excerpt from the Job Content Questionnaire
Figure II.1: Odds of Overgrading by Average GS Grade Across 37 Occupations
Figure II.2: Odds of Undergrading by Average GS Grade Across 37 Occupations

Abbreviations
CPDF Central Personnel Data File
FES Factor Evaluation System
GS General Schedule
MSPB Merit System Protection Board
OPM Office of Personnel Management
PATCO Professional, Administrative, Technical, Clerical, Other
Appendix I

Questionnaire Development, Sample Selection, Response Rate Calculation, and Validation Methodology

Using the competitive bid process, we contracted with a national management consultant—Barrett & Associates, Inc., of Akron, Ohio—to (1) develop a job content questionnaire that enabled the contractor to estimate GS grades for positions in our nationwide sample using input from the incumbents and their supervisors, (2) distribute the questionnaire to a sample of incumbent/supervisor pairs attaining at least an 80 percent response rate, (3) develop a semistructured interview guide for completing desk audits, (4) validate the questionnaire through the use of desk audits, and (5) analyze the questionnaire responses.¹ The analyses were done to determine whether a link existed between the GS grades assigned to positions through the use of the Factor Evaluation System (FES) and the number of women or minorities in occupations that we included in our study.

We selected the sample of incumbent/supervisor pairs and worked closely with the contractor providing supervision throughout the process. Because the preliminary analyses showed that more variation existed within groups of occupations rather than between those groups, we refined the analyses of the questionnaire responses as described in appendix II.

In addition to working with the contractor, we consulted periodically with an ad hoc panel of experts that provided technical guidance during the early design phase of our study. Our panel consisted of

• Ms. Ruth Rogers, former Chief, Standards Division, Department of Personnel, Government of the District of Columbia, who provided staff assistance to the District’s pay equity study;²
• Dr. Donald Schwab, Professor of Business Research and Industrial Relations, University of Wisconsin-Madison, who has written extensively on the issue of comparable worth; and
• Dr. Ronnie Steinberg, Professor of Sociology and Women’s Studies, Temple University, who has considerable experience with pay equity issues and job content questionnaires.

¹Desk audits consist of face-to-face interviews between trained classifiers and the incumbent and supervisor to identify and observe work operations and products. The validation was done to determine whether our questionnaire could be used to estimate GS grades similar to those that would be attained in actual desk audits. When we designed our study, we set a goal of achieving a .80 validity coefficient, which measures the consistency between two data sets.

²According to a District of Columbia official, in July 1990, the District contracted with a consulting firm to do a pay equity study, and in April 1994, the contractor provided a final report with recommendations.
We also consulted with Dr. David Rindskopf, Professor, Graduate School and University Center, City University of New York, for assistance with our sample selection methodology.

After we completed our analyses, we shared our results with and obtained informal comments from five federal officials with classification experience. In addition to a representative from OPM and GAO’s personnel office, we selected three of these experts from a list of recommended candidates provided at our request by the Classification and Compensation Society, a professional organization of federal classifiers and other personnel specialists. The classification experts included a mix of male, female, minority, and nonminority individuals who worked in OPM’s then Personnel Systems and Oversight Group and in the personnel offices of the Federal Aviation Administration, the Department of Defense, the Library of Congress, and the General Accounting Office.

Questionnaire Development and GS Grade Estimation

We developed a job content questionnaire that enabled us to estimate GS grades for positions covered by FES, a point factor position classification system the federal government used to classify 77 of the 455 white-collar occupations when we designed our study.\(^3\) On the basis of the factor descriptions in the FES primary standard, we constructed questions that allowed us to determine the appropriate level for each factor.\(^4\) Table I.1 shows the points associated with each factor level.

---

\(^3\)A point factor position classification system uses a set of factors and factor weights to order positions hierarchically in terms of their value to an employer. FES uses the following nine factors to describe the duties and responsibilities of a position: knowledge required by the position, supervisory controls, guidelines, complexity, scope and effect, personal contacts, purpose of contacts, physical demands, and work environment. Each factor is broken down into graduated levels. Factors are composed of from three to nine levels.

\(^4\)The FES primary standard serves as the framework for individual classification standards written for each occupation. We compared the occupation-specific standards with the primary standard for four occupational series—computer operations, border patrol agent, equal employment opportunity, and equal opportunity compliance—and determined that these occupation-specific standards appeared to be within the framework of the primary standard.
Table I.1: FES Factors and Points by Factor Levels

<table>
<thead>
<tr>
<th>Factor</th>
<th>Levela</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Knowledge required by the position</td>
<td>50</td>
</tr>
<tr>
<td>Supervisory controls</td>
<td>25</td>
</tr>
<tr>
<td>Guidelines</td>
<td>25</td>
</tr>
<tr>
<td>Complexity</td>
<td>25</td>
</tr>
<tr>
<td>Scope and effect</td>
<td>25</td>
</tr>
<tr>
<td>Personal contacts</td>
<td>10</td>
</tr>
<tr>
<td>Purpose of contacts</td>
<td>20</td>
</tr>
<tr>
<td>Physical demands</td>
<td>5</td>
</tr>
<tr>
<td>Work environment</td>
<td>5</td>
</tr>
</tbody>
</table>

aThe blank spaces in columns indicate that no further levels exist for the corresponding factor.

Source: FES Primary Standard.

Figures I.1 and I.2 show an excerpt from the FES primary standard and the resulting question that we included in our job content questionnaire, respectively.
FACTOR 1. KNOWLEDGE REQUIRED BY THE POSITION

Factor 1 measures the nature and extent of information or facts which the workers must understand to do acceptable work (e.g., steps, procedures, practices, rules, policies, theories, principles, and concepts) and the nature and extent of the skills needed to apply those knowledges. To be used as a basis for selecting a level under this factor, a knowledge must be required and applied.

**Level 1-1**

Knowledge of simple, routine, or repetitive tasks or operations which typically includes following step-by-step instructions and requires little or no previous training or experience;

**Level 1-2**

Knowledge of basic or commonly-used rules, procedures, or operations which typically requires some previous training or experience;

**Level 1-8**

Mastery of a professional or administrative field to:

- Apply experimental theories and new developments to problems not susceptible to treatment by accepted methods;

**Level 1-9**

Mastery of a professional field to generate and develop new hypotheses and theories;
### Figure I.2: Excerpt From the Job Content Questionnaire

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>13. To perform your job in a satisfactory manner, what level of knowledge and skill is needed? (Check one.)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>☐ Tasks or operations involving specific step-by-step instructions requiring little or no previous training or experience</td>
</tr>
<tr>
<td>2.</td>
<td>☐ A few basic or commonly used rules, procedures, or operations plus some previous or on-the-job training or experience</td>
</tr>
<tr>
<td>8.</td>
<td>☐ Mastery of a professional, technical or administrative field to the point where one can apply new hypotheses, theories, or applications to applied problems</td>
</tr>
<tr>
<td>9.</td>
<td>☐ Mastery of a professional, technical or administrative field to the point where one can generate and develop new hypotheses and theories</td>
</tr>
</tbody>
</table>

The GS grade is determined by totaling the points assigned to each of the nine factors and using the grade conversion table shown in table I.2.
To determine whether incumbents could readily understand the questionnaire and complete it within a reasonable time period, we completed three sets of pretests. In total, we selected about 30 incumbents and supervisors who were located in the Cleveland metropolitan area and who represented a range of GS grades and occupations covered by FES. After observing them as they completed the questionnaire, we interviewed each incumbent or supervisor to identify how the questionnaire could be improved; we rewrote or edited most of the questions on the basis of the information they provided. During the initial pretest, we also revised our semistructured interview guide for use in completing subsequent desk audits.

On the basis of questionnaire responses received from incumbents and their supervisors, we estimated the appropriate GS grade for each position in our sample by (1) determining the appropriate level (and corresponding points) for each of the nine factors, (2) averaging the total points computed separately for the incumbent and the supervisor, and (3) using the FES points-to-grade conversion table to assign a GS grade to the position. We decided to average the input from the incumbent and the supervisor to balance the views of those who would place more reliance
on the input of incumbents versus that of supervisors. Table I.3 shows an example of how we estimated the GS grade for one position in our sample.

### Table I.3: Example of GS Grade Estimation Based on Incumbent and Supervisor Questionnaire Responses

<table>
<thead>
<tr>
<th>Factor</th>
<th>Incumbent Factor level</th>
<th>Incumbent Points</th>
<th>Supervisor Factor level</th>
<th>Supervisor Points</th>
<th>Average total points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge required by the position</td>
<td>4</td>
<td>550</td>
<td>3</td>
<td>350</td>
<td></td>
</tr>
<tr>
<td>Supervisory controls</td>
<td>2</td>
<td>125</td>
<td>2</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>Guidelines</td>
<td>2</td>
<td>125</td>
<td>1</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Complexity</td>
<td>3</td>
<td>150</td>
<td>1</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Scope and effect</td>
<td>2</td>
<td>75</td>
<td>4</td>
<td>225</td>
<td></td>
</tr>
<tr>
<td>Personal contacts</td>
<td>3</td>
<td>60</td>
<td>1</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Purpose of contacts</td>
<td>1</td>
<td>20</td>
<td>1</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Physical demands</td>
<td>2</td>
<td>20</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Work environment</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td><strong>Total points</strong></td>
<td><strong>1,130</strong></td>
<td><strong>805</strong></td>
<td><strong>967</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated GS grade</td>
<td></td>
<td></td>
<td></td>
<td>GS-5</td>
<td></td>
</tr>
</tbody>
</table>

Note: The blank spaces in columns indicate that we estimated GS grades by averaging the total points on the basis of input from the incumbent and the supervisor rather than the points assigned to each factor.

### Sample Selection

We selected our nationwide sample from the full-time permanent nonsupervisory incumbents in the 77 occupations for which an FES standard had been in existence for at least 1 year when we completed our study design in May 1992. Because of limited resources, we excluded incumbents who (1) were stationed outside of the continental United States or (2) worked for agencies with less than 500 employees. Each of the 77 occupations included from 35 to 91,769 incumbents for a total of 441,189 full-time permanent nonsupervisory incumbents. Women constituted about 61 percent of the total workforce covered by FES and minorities, approximately 32 percent. We used OPM’s Central Personnel Data File (CPDF), updated as of June 1992, to select our sample. These were the most recent data available when we designed our study. We did not independently verify the accuracy of this database.

We defined each of the 77 occupations covered by FES as having a high, medium, or low representation of women or minorities. On the basis of the current literature, we defined occupations with high, medium, and low
female representation as those in which women represented 70 percent or more, 31 to 69 percent, and 30 percent or less of the incumbents, respectively. The literature contained no consistent basis on which to define minority representation; therefore, on the basis of our past work, we initially adopted a standard of more than 48 percent (150 percent of the 32 percent minority workforce representation) to define high minority representation; less than 16 percent (50 percent of 32 percent), low minority representation; and 16 to 48 percent, medium minority representation. However, in order to include in our sample at least two occupations with each possible mix of gender and minority representation (e.g., low female, high minority representation), we found it necessary to adjust our definitions for occupations with high, medium, and low minority representation to those in which minorities represented more than 41 percent, 23 to 41 percent, and less than 23 percent of incumbents, respectively. Table I.4 shows the distribution of the 77 occupations in a nine-cell matrix configured according to female and minority representation.

Table I.4: Distribution of the 77 FES Occupations

<table>
<thead>
<tr>
<th>Female representation</th>
<th>Minority representation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>High</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>Medium</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Low</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>27</td>
</tr>
</tbody>
</table>

Source: GAO analysis of data obtained from OPM.

From the 77 occupations covered by FES, we selected 58 on the basis of four characteristics—number of incumbents, PATCO (professional,
administrative, technical, clerical, and other) category,\(^5\) job family,\(^6\) and GS grade distribution. First, we chose occupations in each matrix cell with the largest number of incumbents. Second, we examined the PATCO category of the occupations selected within each cell and chose additional occupations to include all categories. Third, we reviewed the job families represented by the occupations already selected and chose additional occupations increasing the number of families included in our study to 17 out of a total of 22. Finally, we examined the GS grade distributions of the occupations selected and chose occupations that included grades not previously selected within each row and column of the matrix. Table I.5 shows the female and minority representation, PATCO category, and number of full-time permanent nonsupervisory incumbents for each of the 77 occupations covered by FES when we selected our sample as well as the 58 occupations selected for inclusion in our sample.

\(^5\) OPM assigns a PATCO category to each white-collar occupation on the basis of the general subject matter of work, level of difficulty or responsibility, and educational requirements.

Professional occupations require incumbents to use discretion and judgment to apply knowledge acquired through education or training equivalent to a bachelor's degree in a specialized field.

Administrative occupations involve the exercise of analytical ability and personal responsibility to apply concepts and practices—typically learned through a general college education or progressively responsible work—to one or more fields of management.

Technical occupations consist of nonroutine work that is learned on-the-job or from specialized training less than that represented by college graduation to support professional or administrative fields.

Clerical occupations require incumbents to do structured work according to established policies, which are learned through training or work experience, to support office operations.

Other occupations include those miscellaneous occupations that are not included in one of the four other categories.

\(^6\) Job families consist of related occupations that are numbered using the same multiple of 100. For example, GS-300, the general administration, clerical, and office services group includes the mail and file (GS-305), computer operation (GS-332), and telecommunications (GS-391) occupations.
Table I.5: Female and Minority Representation, PATCO Category, and Number of Incumbents for the 77 Occupations Covered by FES

<table>
<thead>
<tr>
<th>Occupational series a</th>
<th>Representation</th>
<th>PATCO b</th>
<th>Full-time permanent nonsupervisory incumbents</th>
</tr>
</thead>
<tbody>
<tr>
<td>318</td>
<td>Secretary</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>334</td>
<td>Computer specialist</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>1102</td>
<td>Contracting</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>525</td>
<td>Accounting technician</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>621</td>
<td>Nursing assistant</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>305</td>
<td>Mail and file</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>592</td>
<td>Tax examining</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>1910</td>
<td>Quality assurance</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>203</td>
<td>Personnel clerical and assistance</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>511</td>
<td>Auditing</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>560</td>
<td>Budget analysis</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>830</td>
<td>Mechanical engineering</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>679</td>
<td>Medical clerk</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>510</td>
<td>Accounting</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>335</td>
<td>Computer clerk and assistant</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>332</td>
<td>Computer operation</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>610</td>
<td>Nurse</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>1320</td>
<td>Chemistry</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>83</td>
<td>Police</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>80</td>
<td>Security administration</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>391</td>
<td>Telecommunications</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>644</td>
<td>Medical technologist</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>457</td>
<td>Soil conservation</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>540</td>
<td>Voucher examining</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>(561)</td>
<td>(Budget clerical and assistance)</td>
<td>(High)</td>
<td>(Medium)</td>
</tr>
<tr>
<td>85</td>
<td>Security guard</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>(1370)</td>
<td>(Cartography)</td>
<td>(Low)</td>
<td>(Low)</td>
</tr>
<tr>
<td>460</td>
<td>Forestry</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>(544)</td>
<td>(Civilian pay)</td>
<td>(High)</td>
<td>(Medium)</td>
</tr>
<tr>
<td>(819)</td>
<td>(Environmental engineering)</td>
<td>(Low)</td>
<td>(Low)</td>
</tr>
<tr>
<td>950</td>
<td>Paralegal specialist</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>18</td>
<td>Safety and occupational health management</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>1035</td>
<td>Public affairs</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>1896</td>
<td>Border patrol agent</td>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

(continued)
### Appendix I
Questionnaire Development, Sample
Selection, Response Rate Calculation, and
Validation Methodology

<table>
<thead>
<tr>
<th>Occupational series</th>
<th>Representation</th>
<th>PATCO</th>
<th>Full-time permanent nonsupervisory incumbents</th>
</tr>
</thead>
<tbody>
<tr>
<td>475</td>
<td>Agricultural management</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>1311</td>
<td>Physical science technician</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>(545)</td>
<td>(Military pay)</td>
<td>(High)</td>
<td>(Medium)</td>
</tr>
<tr>
<td>(840)</td>
<td>(Nuclear engineering)</td>
<td>(Low)</td>
<td>(Low)</td>
</tr>
<tr>
<td>681</td>
<td>Dental assistant</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>620</td>
<td>Practical nurse</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>647</td>
<td>Diagnostic radiologic technologist</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>2134</td>
<td>Shipment clerical and assistance</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>622</td>
<td>Medical supply aide and technician</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>260</td>
<td>Equal employment opportunity</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>(486)</td>
<td>(Wildlife biology)</td>
<td>(Low)</td>
<td>(Low)</td>
</tr>
<tr>
<td>1173</td>
<td>Housing management</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>(458)</td>
<td>(Soil conservation technician)</td>
<td>(Medium)</td>
<td>(Low)</td>
</tr>
<tr>
<td>649</td>
<td>Medical instrument technician</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>350</td>
<td>Equipment operator</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>808</td>
<td>Architecture</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>360</td>
<td>Equal opportunity compliance</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>(1980)</td>
<td>(Agricultural commodity grading)</td>
<td>(Low)</td>
<td>(Low)</td>
</tr>
<tr>
<td>1371</td>
<td>Cartographic technician</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>(482)</td>
<td>(Fishery biology)</td>
<td>(Low)</td>
<td>(Low)</td>
</tr>
<tr>
<td>690</td>
<td>Industrial hygiene</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>1822</td>
<td>Mine safety and health</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>1889</td>
<td>Import specialist</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>630</td>
<td>Dietician and nutritionist</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>436</td>
<td>Plant protection and quarantine</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>701</td>
<td>Veterinary medical science</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>188</td>
<td>Recreation specialist</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>1040</td>
<td>Language specialist</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>(651)</td>
<td>(Respiratory therapist)</td>
<td>(Medium)</td>
<td>(High)</td>
</tr>
<tr>
<td>361</td>
<td>Equal opportunity assistance</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>270</td>
<td>Federal retirement benefits</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>1720</td>
<td>Education program</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>(485)</td>
<td>(Wildlife refuge management)</td>
<td>(Low)</td>
<td>(Low)</td>
</tr>
<tr>
<td>682</td>
<td>Dental hygiene</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>(2121)</td>
<td>(Railroad safety)</td>
<td>(Low)</td>
<td>(Low)</td>
</tr>
<tr>
<td>309</td>
<td>Correspondence clerk</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>(2123)</td>
<td>(Motor carrier safety)</td>
<td>(Low)</td>
<td>(Low)</td>
</tr>
</tbody>
</table>

(continued)
### Appendix I
Questionnaire Development, Sample Selection, Response Rate Calculation, and Validation Methodology

<table>
<thead>
<tr>
<th>Occupational series</th>
<th>Representation</th>
<th>Full-time permanent nonsupervisory incumbents</th>
</tr>
</thead>
<tbody>
<tr>
<td>(669) (Medical records administration)</td>
<td>(High) (Medium) (A)</td>
<td>(170)</td>
</tr>
<tr>
<td>(1161) (Crop insurance administration)</td>
<td>(Low) (Low) (A)</td>
<td>(109)</td>
</tr>
<tr>
<td>(1162) (Crop insurance underwriting)</td>
<td>(Low) (Low) (A)</td>
<td>(54)</td>
</tr>
<tr>
<td>631 Occupational therapist</td>
<td>High Low P</td>
<td>49</td>
</tr>
<tr>
<td>(1884) (Customs patrol officer)</td>
<td>(Low) (High) (O)</td>
<td>(47)</td>
</tr>
<tr>
<td>(160) (Civil rights analysis)</td>
<td>(Medium) (High) (A)</td>
<td>(35)</td>
</tr>
</tbody>
</table>

**Subtotal/included in study**  
412,279

**Subtotal/not included in study**  
(28,910)

**Total**  
441,189

Note: This table includes occupational series, or occupations, covered by FES that had been in existence for at least 1 year as of May 1992.

*We selected incumbents for our sample from the 58 occupational series not shown in parentheses.

*The PATCO category indicates the general characteristics of the work done within each federal white-collar occupation and is represented as Professional, Administrative, Technical, Clerical, or Other.

*The PATCO category for this occupation is Clerical for GS grades 1 to 3 and Technical for grades 4 and above.

*The PATCO category for this occupation is Clerical for GS grades 1 to 5 and Technical for grades 6 and above.

*The PATCO category for this occupation is Clerical for GS grades 1 to 4 and Technical for grades 5 and above.

Source: GAO analysis of data obtained from OPM.

We selected no more than 11 occupations for each matrix cell because we planned to validate the job content questionnaire by completing two sets of 100 desk audits that would each (1) include at least one position from each occupation in our sample and (2) be evenly distributed among the 9 matrix cells. Table I.6 shows the distribution of the 58 occupations selected for inclusion in our sample by matrix cell.
Table I.6: Distribution of the 58 Occupations Included in the Sample

<table>
<thead>
<tr>
<th>Female representation</th>
<th>Minority representation</th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>High</td>
<td>6</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>High</td>
<td>Medium</td>
<td>10</td>
<td>10</td>
<td>24</td>
</tr>
<tr>
<td>High</td>
<td>Low</td>
<td>2</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>18</td>
<td>23</td>
<td>58</td>
</tr>
</tbody>
</table>

Source: GAO analysis of data obtained from OPM.

After selecting the 58 occupations, we completed a pilot test by distributing the job content questionnaire to 389 pairs of incumbents and their supervisors who were located in either the Cleveland, OH, or Washington, D.C. areas and who represented a range of GS grades and occupations included in our sample. Because the CPDF does not identify a specific address or supervisor for incumbents, we forwarded the questionnaires to the appropriate agency personnel offices for distribution. Pairs of trained job analysts composed of at least one female and one minority completed the first set of desk audits for 100 of the 257 positions for which the incumbent and the supervisor returned a questionnaire. Of the 100 positions, we included in our analyses the 84 positions for which the incumbent and the supervisor provided complete responses to all questionnaire items. To ensure the independence of the validation process, the contractor instructed the job analysts not to review the incumbent/supervisor responses to the questionnaire for any position before they completed a desk audit.

On the basis of the pilot test, we determined that (1) incumbents at higher GS grades tended to undervalue their positions when compared to the desk audit, while incumbents at lower grades tended to overvalue their positions and (2) the grades of incumbents included in our study were not evenly distributed across the matrix rows and columns. As a result, the validity coefficient between GS grades estimated on the basis of the desk audits versus incumbent/supervisor questionnaire responses was .74. Because the combination of the two effects threatened to distort the results of our planned statistical comparisons, we stratified our sample by classifying the incumbents of the 58 occupations into one of seven groups, or strata, on the basis of GS grade: grades 1 to 4, grade 5, grade 6, grades 7 to 8, grades 9 to 10, grade 11, and grades 12 to 15. We then determined the

---

7Although the majority of the job analysts had extensive desk audit experience, the contractor provided the analysts with comprehensive training that included writing job descriptions and classifying jobs through the use of FES.
number of incumbent/supervisor pairs needed within each matrix cell using a complex balancing design to ensure that the row and column totals would be about equal. This stratified sampling strategy enabled us to balance the grade distribution of incumbents selected within each row and column of the matrix, and thus, eliminate any grade level effect in subsequent analyses.

We randomly selected our sample of full-time permanent nonsupervisory incumbents from the 58 occupations that met the criteria for each matrix cell and GS grade stratum and distributed the job content questionnaire to 2,233 incumbents and their supervisors. Because of the random selection process, we did not select incumbents from 3 of the 58 occupational series—import specialist, education program, and correspondence clerk. Table 1.7 shows the distribution of the 2,233 incumbents in a matrix configured according to female and minority representation and strata.
Table I.7: Distribution of the 2,233 Incumbents by Female and Minority Representation and Strata

<table>
<thead>
<tr>
<th>GS grade strata</th>
<th>High</th>
<th>Medium</th>
<th>Low</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS-1 to 4</td>
<td>45</td>
<td>44</td>
<td>1</td>
<td>90</td>
</tr>
<tr>
<td>GS-5</td>
<td>65</td>
<td>66</td>
<td>19</td>
<td>150</td>
</tr>
<tr>
<td>GS-6</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>90</td>
</tr>
<tr>
<td>GS-7 to 8</td>
<td>30</td>
<td>30</td>
<td>29</td>
<td>89</td>
</tr>
<tr>
<td>GS-9 to 10</td>
<td>30</td>
<td>30</td>
<td>12</td>
<td>72</td>
</tr>
<tr>
<td>GS-11</td>
<td>3</td>
<td>44</td>
<td>15</td>
<td>62</td>
</tr>
<tr>
<td>GS-12 to 15</td>
<td>1</td>
<td>144</td>
<td>1</td>
<td>146</td>
</tr>
<tr>
<td><strong>Subtotal</strong></td>
<td>204</td>
<td>388</td>
<td>107</td>
<td>699</td>
</tr>
<tr>
<td>GS-1 to 4</td>
<td>30</td>
<td>16</td>
<td>44</td>
<td>90</td>
</tr>
<tr>
<td>GS-5</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>150</td>
</tr>
<tr>
<td>GS-6</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>90</td>
</tr>
<tr>
<td>GS-7 to 8</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>90</td>
</tr>
<tr>
<td>GS-9 to 10</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>90</td>
</tr>
<tr>
<td>GS-11</td>
<td>42</td>
<td>28</td>
<td>20</td>
<td>90</td>
</tr>
<tr>
<td>GS-12 to 15</td>
<td>115</td>
<td>19</td>
<td>50</td>
<td>184</td>
</tr>
<tr>
<td><strong>Subtotal</strong></td>
<td>327</td>
<td>203</td>
<td>254</td>
<td>784</td>
</tr>
<tr>
<td>GS-1 to 4</td>
<td>15</td>
<td>30</td>
<td>45</td>
<td>90</td>
</tr>
<tr>
<td>GS-5</td>
<td>35</td>
<td>34</td>
<td>81</td>
<td>150</td>
</tr>
<tr>
<td>GS-6</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>90</td>
</tr>
<tr>
<td>GS-7 to 8</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>90</td>
</tr>
<tr>
<td>GS-9 to 10</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>90</td>
</tr>
<tr>
<td>GS-11</td>
<td>45</td>
<td>18</td>
<td>27</td>
<td>90</td>
</tr>
<tr>
<td>GS-12 to 15</td>
<td>34</td>
<td>19</td>
<td>97</td>
<td>150</td>
</tr>
<tr>
<td><strong>Subtotal</strong></td>
<td>219</td>
<td>191</td>
<td>340</td>
<td>750</td>
</tr>
</tbody>
</table>

| **Total**       | 750  | 782    | 701 | 2,233 |

Source: GAO analysis of data obtained from OPM.

In addition to the sample of 2,233 incumbent and supervisor pairs, we also selected a supplemental sample of 303 pairs to enable us to complete a second set of desk audits expeditiously. We selected the supplemental sample from those incumbents in the 58 occupations who were located in one of three geographical areas—Washington, D.C., Dayton/Cincinnati, OH, or Los Angeles, CA—to represent incumbents working in the eastern, central, and western United States. Because we did not randomly select the pairs in the supplemental sample from all incumbents working in the
Appendix I
Questionnaire Development, Sample
Selection, Response Rate Calculation, and
Validation Methodology

58 occupations covered by FES, we did not include their questionnaire responses in our response rate calculation or analyses.

Response Rate Calculation

We forwarded the job content questionnaires to the appropriate agency personnel offices for distribution to each of the 2,233 incumbent/supervisor pairs. For 179 pairs, agency officials notified us that either the incumbent or the supervisor no longer held the position indicated by the CPDF or were located outside of the United States. As planned, we eliminated these pairs from our study. Of the remaining 2,054 pairs, we received questionnaires from 1,633 pairs of respondents, for a response rate of 80 percent. Table I.8 shows the disposition of each of the 2,233 incumbent/supervisor pairs in our sample.

<table>
<thead>
<tr>
<th>Incumbent/supervisor pairs eliminated from the study</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Incumbent resigned or terminated</td>
<td>63</td>
</tr>
<tr>
<td>Incumbent transferred to another agency</td>
<td>48</td>
</tr>
<tr>
<td>Incumbent in different occupation than one identified by the CPDF</td>
<td>29</td>
</tr>
<tr>
<td>Incumbent or supervisor retired</td>
<td>26</td>
</tr>
<tr>
<td>Incumbent or supervisor outside the United States</td>
<td>3</td>
</tr>
<tr>
<td>Incumbent in position not covered by FES</td>
<td>9</td>
</tr>
<tr>
<td>Incumbent deceased</td>
<td>1</td>
</tr>
</tbody>
</table>

**Subtotal** 179

<table>
<thead>
<tr>
<th>Incumbent/supervisor pairs in the study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respondents</td>
</tr>
<tr>
<td>Nonrespondents</td>
</tr>
</tbody>
</table>

**Subtotal** 2,054

**Total incumbent/supervisor pairs in sample** 2,233

*Of the 179 incumbent/supervisor pairs, we eliminated 177 pairs from the study on the basis of factors relating to the incumbent. Two were eliminated on the basis of factors relating to the supervisor.

*The 421 pairs of nonrespondents consisted of (1) 349 pairs for which either the incumbent, the supervisor, or both did not respond to our questionnaire after three follow-up efforts and (2) 72 pairs for which we were unable to forward a questionnaire to either the incumbent, the supervisor, or both for reasons which did not eliminate the pairs from our study (e.g., either the incumbent or the supervisor was on extended leave).
For 201 of the 1,633 pairs of respondents, either the incumbent, the supervisor, or both left some questionnaire items unanswered. Therefore, we included the remaining 1,432 pairs in our preliminary analyses. On the basis of the incumbent/supervisor questionnaire responses, we estimated the GS grade for each of the 1,432 positions and compared this grade with the position’s actual grade. We considered positions to be appropriately graded when the actual grade equalled the questionnaire grade; otherwise, we considered the position to be overgraded if the actual grade was higher than the questionnaire grade; and undergraded if the actual grade was lower than the questionnaire grade.

When we designed this study, we planned to compare the overgrading and undergrading among groups of occupations with different gender and minority representations or among matrix rows and columns. However, preliminary analyses showed that more variation in overgrading and undergrading existed within the same row or column rather than between different rows or columns. For this reason, we analyzed the 37 occupations for which we had received completed questionnaires from at least 10 or more incumbent/supervisor pairs. We did not include the remaining 18 occupations in our analyses because we could not obtain reliable estimates of overgrading and undergrading for an occupation with less than 10 pairs of respondents. The 37 occupations included 1,358 incumbent/supervisor pairs and represented about 79 percent of the incumbents covered by FES. Table I.9 shows the gender and minority representation of the 37 occupations included in our study.

The 18 occupational series for which we received completed questionnaires from less than 10 incumbent/supervisor pairs included tax examining, computer clerk and assistant, security administration, medical technologist, voucher examining, practical nurse, shipment clerical and assistance, medical supply aide and technician, medical instrument technician, equipment operator, architecture, mine safety and health, dietician and nutritionist, veterinary medical science, recreation specialist, language specialist, equal opportunity assistance, and federal retirement benefits.
### Table I.9: Female and Minority Representation of the 37 Occupations Included in the Study

<table>
<thead>
<tr>
<th>Female representation</th>
<th>Minority representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Dental assistant</td>
<td>Accounting technician</td>
</tr>
<tr>
<td>Medical clerk</td>
<td>Budget analysis</td>
</tr>
<tr>
<td>Personnel clerical and assistance</td>
<td>Nurse</td>
</tr>
<tr>
<td></td>
<td>Paralegal specialist</td>
</tr>
<tr>
<td></td>
<td>Secretary</td>
</tr>
<tr>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Computer operation</td>
<td>Accounting</td>
</tr>
<tr>
<td>Equal employment opportunity</td>
<td>Computer specialist</td>
</tr>
<tr>
<td>Equal opportunity compliance</td>
<td>Contracting</td>
</tr>
<tr>
<td>Mail and file</td>
<td>Diagnostic radiologic technologist</td>
</tr>
<tr>
<td>Nursing assistant</td>
<td>Housing management</td>
</tr>
<tr>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Border patrol agent</td>
<td>Plant protection and quarantine</td>
</tr>
<tr>
<td>Security guard</td>
<td>Police</td>
</tr>
<tr>
<td></td>
<td>Chemistry</td>
</tr>
<tr>
<td></td>
<td>Forestry</td>
</tr>
<tr>
<td></td>
<td>Mechanical engineering</td>
</tr>
<tr>
<td></td>
<td>Quality assurance</td>
</tr>
<tr>
<td></td>
<td>Physical science technician</td>
</tr>
<tr>
<td></td>
<td>Safety and occupational health management</td>
</tr>
<tr>
<td></td>
<td>Soil conservation</td>
</tr>
<tr>
<td></td>
<td>Telecommunications</td>
</tr>
</tbody>
</table>

Source: GAO analysis of data obtained from OPM.

**Questionnaire Validation**

We validated our job content questionnaire by using a second set of desk audits. To simplify selecting positions that represented a broad range of grades, we combined our seven strata into three groups: high (GS-11 to 15), medium (GS-6 to 10), and low (GS-1 to 5). We then scheduled interviews with incumbent/supervisor pairs, which allowed us to complete desk...
audits that were distributed evenly across the three GS grade ranges, the three selected geographical locations, and the nine matrix cells. Again, to ensure the independence of the validation process, the contractor instructed the job analysts not to review the incumbent/supervisor responses to the questionnaire for any position before they completed a desk audit.

Although we completed 100 desk audits, we validated the questionnaire on the basis of the 78 positions that represented the 37 occupations included in our analyses and for which the incumbent and supervisor provided complete responses to the questionnaire. For the 78 positions, we completed 24 desk audits for high, 35 for medium, and 19 for low graded positions; and 37 desk audits in Washington, D.C.; 21 in Los Angeles, CA; and 20 in Dayton/Cincinnati, OH. Table I.10 shows the distribution of the 78 positions for which a desk audit was completed by occupation and matrix cell.

<p>| Table I.10: Distribution of Desk Audits Completed According to Female and Minority Representation for the 37 Occupations Included in the Study |
|-------------------------------------------------|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Male representation</th>
<th>High</th>
<th>Medium</th>
<th>Low</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dental assistant</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Medical clerk</td>
<td>3</td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Personnel clerical and assistance</td>
<td>6</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Subtotal</td>
<td>9</td>
<td></td>
<td>6</td>
<td>24</td>
</tr>
</tbody>
</table>

(continued)
### Male representation

| Table I.11: Validity Coefficients Between GS Grades Estimated on the Basis of 78 Desk Audits vs. Incumbent/Supervisor Questionnaire Responses
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Minority representation</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
<td>Total</td>
</tr>
<tr>
<td>Computer operation</td>
<td>2</td>
<td>Accounting</td>
<td>2</td>
<td>Auditing</td>
</tr>
<tr>
<td>Equal employment opportunity</td>
<td>2</td>
<td>Computer specialist</td>
<td>3</td>
<td>Cartographic technician</td>
</tr>
<tr>
<td>Equal opportunity compliance</td>
<td>0</td>
<td>Contracting</td>
<td>3</td>
<td>Industrial hygiene</td>
</tr>
<tr>
<td>Mail and file</td>
<td>2</td>
<td>Diagnostic radiologic technologist</td>
<td>1</td>
<td>Public affairs</td>
</tr>
<tr>
<td>Nursing assistant</td>
<td>1</td>
<td>Housing management</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Subtotal</td>
<td>7</td>
<td>Subtotal</td>
<td>10</td>
<td>Subtotal</td>
</tr>
<tr>
<td>Border patrol agent</td>
<td>4</td>
<td>Plant protection and quarantine</td>
<td>0</td>
<td>Agricultural management</td>
</tr>
<tr>
<td>Security guard</td>
<td>3</td>
<td>Police</td>
<td>6</td>
<td>Chemistry</td>
</tr>
<tr>
<td>Forestry</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mechanical engineering</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality assurance</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical science technician</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safety and occupational health management</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soil conservation</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telecommunications</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal</td>
<td>7</td>
<td>Subtotal</td>
<td>6</td>
<td>Subtotal</td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>22</td>
<td>33</td>
<td>78</td>
</tr>
</tbody>
</table>

The resulting validity coefficient between GS grades estimated on the basis of the 78 desk audits versus incumbent/supervisor questionnaire responses was .80, thus meeting the goal we set when we designed the study. This is the only classification study we are aware of in which desk audits were used to validate the questionnaire results. Table I.11 shows the validity coefficients between GS grades estimated on the basis of desk audits.
versus the questionnaire responses for incumbent/supervisor pairs, supervisors, and incumbents.

<table>
<thead>
<tr>
<th>Source of questionnaire responses</th>
<th>Validity coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incumbent/supervisor pairs</td>
<td>.80</td>
</tr>
<tr>
<td>Supervisors</td>
<td>.75</td>
</tr>
<tr>
<td>Incumbents</td>
<td>.73</td>
</tr>
</tbody>
</table>

Note: While we validated our questionnaire on the basis of the responses from incumbent/supervisor pairs, we reported the validity coefficients for supervisors only and incumbents only for the purpose of full disclosure.

For the 37 occupations included in our analyses, we used loglinear statistical techniques to analyze the odds of individual occupations being overgraded or undergraded versus appropriately graded (1) in relation to GS grades and (2) when the female or minority representation of an occupation was high, medium, or low. The strength of this particular statistical approach was that multiple variables could be analyzed simultaneously, thereby enabling us to examine complex relationships in the data. Appendix II provides more detailed information on our loglinear methodology, the loglinear models tested, and the results obtained.
This appendix provides additional technical detail on our analytical approach. It contains a general description of loglinear methodology, describes the variables analyzed, and presents the loglinear models tested and the results obtained in each analysis.

Data Analysis Approach

We used loglinear analyses to examine the relationship between the likelihood of being overgraded or undergraded and (1) the average GS grade of the sampled incumbents by occupation and (2) female and minority representation. We first looked at the relationship between the likelihood of being overgraded or undergraded and female or minority representation without taking the effect of the average GS grade into consideration. For each analysis, we considered the preferred model to be the simplest one that fit the data and could not be significantly improved by more complex models. The preferred model included those components that had statistically significant relationships with effects after we controlled for the influences of other factors. Hence, the estimates we obtained were net effects determined after the association of each variable with all other variables had been taken into account or statistically eliminated.

On the basis of the preferred model, we estimated the direction and magnitude of the relationships using odds and odds ratios. The odds indicated the likelihood that an outcome would occur given a particular factor or combination of factors, and the odds ratios indicated the size of the effect of the various factors on that likelihood. The more the odds ratio diverges from 1.0, the stronger the relationship.

Occupations Included in Our Study

All of our analyses were based on the 37 occupations shown in table II.1. The table also indicates by occupation whether the female or minority representation was high, medium, or low; the average GS grade of the incumbents in the sample; whether the incumbents were overgraded, aptly graded, or undergraded; and the odds of being either overgraded or

---

1We opted to use loglinear models rather than logistic regression models because of the trichotomous nature (i.e., three characteristics as opposed to being of a continuous nature) of the dependent variable. Logistic regression models could have been used had we chosen to look separately at the two odds that we examined, i.e., the odds of overgrading versus aptly grading and the odds of undergrading versus aptly grading. This would have allowed us to control simultaneously for the average GS grade, gender, and minority status had we wanted to determine how the relationship between job content and GS grades assigned using FES varied on the basis of the gender and minority status of incumbents. Because our objective was to determine whether the relationship varied on the basis of the proportion of women and minorities in occupations, we chose the simpler modeling strategy.
undergraded versus aptly graded. For each occupation or group of occupations, we derived the odds of being overgraded versus aptly graded by dividing the number of incumbents that were overgraded by the number aptly graded. The odds of being undergraded versus aptly graded were similarly calculated.

We defined occupations with high, medium, and low female representation as those in which women represented 70 percent or more, 31 to 69 percent, and 30 percent or less of incumbents, respectively. We defined occupations with high, medium, and low minority representation as those in which minorities represented more than 41 percent, 22 to 41 percent, or less than 22 percent of incumbents, respectively.

On the basis of the average points derived from the questionnaires completed by the incumbent and the supervisor, we determined the GS grade for each position. We compared that questionnaire grade with the incumbent's actual GS grade and considered the position to be aptly, or appropriately, graded if the questionnaire grade and the actual grade were the same. Otherwise, we considered positions to be overgraded if the actual grade was higher than the questionnaire grade and undergraded if the actual grade was lower than the questionnaire grade.
### Table II.1: Female and Minority Representation, Average GS Grades, Observed Sample Frequencies, and Odds of Being Overgraded and Undergraded Versus Aptly Graded for 37 Occupations Without Controlling for GS Grade

<table>
<thead>
<tr>
<th>Female Representation</th>
<th>Minority Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Subtotal for occupations with high female representation</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Subtotal for occupations with medium female representation</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>
### Appendix II

#### Loglinear Methodology and Analysis Results

Ranks and observed frequencies of various occupations are presented with associated odds ratios. The table includes GS occupational series, average grade in sample, observed frequencies (Over graded, Aptly graded, Under graded), and odds ratios comparing over versus aptly, and under versus aptly.

<table>
<thead>
<tr>
<th>GS occupational series</th>
<th>Average GS grade in sample</th>
<th>Observed frequencies</th>
<th>Odds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Over graded</td>
<td>Aptly graded</td>
</tr>
<tr>
<td>Personnel clerical and assistance</td>
<td>6.81</td>
<td>33</td>
<td>20</td>
</tr>
<tr>
<td>Medical clerk</td>
<td>4.45</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Dental assistant</td>
<td>5.12</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Budget analysis</td>
<td>11.98</td>
<td>84</td>
<td>1</td>
</tr>
<tr>
<td>Secretary</td>
<td>5.40</td>
<td>20</td>
<td>24</td>
</tr>
<tr>
<td>Nurse</td>
<td>11.20</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Paralegal specialist</td>
<td>11.56</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>Accounting technician</td>
<td>5.62</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Dental hygiene</td>
<td>6.56</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Occupational therapist</td>
<td>10.50</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>186</th>
<th>68</th>
<th>142</th>
<th>396</th>
<th>2.72</th>
<th>2.08</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equal opportunity compliance</td>
<td>11.86</td>
<td>46</td>
<td>3</td>
<td>0</td>
<td>49</td>
<td>13.29</td>
</tr>
<tr>
<td>Equal employment opportunity</td>
<td>12.26</td>
<td>41</td>
<td>0</td>
<td>1</td>
<td>42</td>
<td>83.00</td>
</tr>
<tr>
<td>Computer operations</td>
<td>7.93</td>
<td>36</td>
<td>3</td>
<td>2</td>
<td>41</td>
<td>10.43</td>
</tr>
<tr>
<td>Nursing assistant</td>
<td>4.90</td>
<td>8</td>
<td>8</td>
<td>13</td>
<td>29</td>
<td>1.00</td>
</tr>
<tr>
<td>Mail and file</td>
<td>4.20</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>10</td>
<td>0.45</td>
</tr>
<tr>
<td>Diagnostic radiologic technologist</td>
<td>5.94</td>
<td>5</td>
<td>6</td>
<td>23</td>
<td>34</td>
<td>0.85</td>
</tr>
<tr>
<td>Computer specialist</td>
<td>9.67</td>
<td>26</td>
<td>3</td>
<td>4</td>
<td>33</td>
<td>7.57</td>
</tr>
<tr>
<td>Contracting</td>
<td>8.57</td>
<td>17</td>
<td>7</td>
<td>4</td>
<td>28</td>
<td>2.33</td>
</tr>
<tr>
<td>Accounting</td>
<td>9.20</td>
<td>10</td>
<td>2</td>
<td>3</td>
<td>15</td>
<td>4.20</td>
</tr>
<tr>
<td>Housing management</td>
<td>5.45</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>11</td>
<td>0.27</td>
</tr>
<tr>
<td>Cartographic technician</td>
<td>6.04</td>
<td>45</td>
<td>25</td>
<td>29</td>
<td>99</td>
<td>1.78</td>
</tr>
<tr>
<td>Public affairs</td>
<td>10.38</td>
<td>41</td>
<td>5</td>
<td>7</td>
<td>53</td>
<td>7.55</td>
</tr>
<tr>
<td>Auditing</td>
<td>10.74</td>
<td>42</td>
<td>3</td>
<td>1</td>
<td>46</td>
<td>12.14</td>
</tr>
<tr>
<td>Industrial hygiene</td>
<td>9.74</td>
<td>16</td>
<td>4</td>
<td>7</td>
<td>27</td>
<td>3.67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>336</th>
<th>79</th>
<th>102</th>
<th>517</th>
<th>4.23</th>
<th>1.29</th>
</tr>
</thead>
<tbody>
<tr>
<td>Border patrol agent</td>
<td>10.75</td>
<td>85</td>
<td>4</td>
<td>3</td>
<td>92</td>
<td>19.00</td>
</tr>
<tr>
<td>Security guard</td>
<td>5.15</td>
<td>16</td>
<td>9</td>
<td>21</td>
<td>46</td>
<td>1.74</td>
</tr>
<tr>
<td>Police</td>
<td>5.34</td>
<td>4</td>
<td>16</td>
<td>38</td>
<td>58</td>
<td>0.27</td>
</tr>
</tbody>
</table>
### Appendix II
Loglinear Methodology and Analysis Results

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th></th>
<th></th>
<th>Minority</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Med</td>
<td>Low</td>
<td>High</td>
<td>Med</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal</td>
<td>for</td>
<td>occupations</td>
<td>with</td>
<td>low</td>
<td>female</td>
<td>representation</td>
</tr>
<tr>
<td></td>
<td>occupations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Observed frequencies and Odds

<table>
<thead>
<tr>
<th>GS occupational series</th>
<th>Average GS grade in sample</th>
<th>Over graded</th>
<th>Aptly graded</th>
<th>Under graded</th>
<th>Total</th>
<th>Over versus aptly&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Under versus aptly&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>436 Plant protection and quarantine</td>
<td>9.13</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td>16</td>
<td>5.80</td>
<td>0.20</td>
</tr>
<tr>
<td>1311 Physical science technician</td>
<td>5.55</td>
<td>31</td>
<td>30</td>
<td>19</td>
<td>80</td>
<td>1.03</td>
<td>0.64</td>
</tr>
<tr>
<td>830 Mechanical engineering</td>
<td>10.95</td>
<td>34</td>
<td>6</td>
<td>3</td>
<td>43</td>
<td>5.31</td>
<td>0.54</td>
</tr>
<tr>
<td>1910 Quality assurance</td>
<td>9.95</td>
<td>17</td>
<td>2</td>
<td>3</td>
<td>22</td>
<td>7.00</td>
<td>1.40</td>
</tr>
<tr>
<td>460 Forestry</td>
<td>8.25</td>
<td>6</td>
<td>5</td>
<td>9</td>
<td>20</td>
<td>1.18</td>
<td>1.73</td>
</tr>
<tr>
<td>1320 Chemistry</td>
<td>10.88</td>
<td>13</td>
<td>2</td>
<td>2</td>
<td>17</td>
<td>5.40</td>
<td>1.00</td>
</tr>
<tr>
<td>457 Soil conservation</td>
<td>8.75</td>
<td>10</td>
<td>3</td>
<td>3</td>
<td>16</td>
<td>3.00</td>
<td>1.00</td>
</tr>
<tr>
<td>475 Agricultural management</td>
<td>9.21</td>
<td>9</td>
<td>3</td>
<td>2</td>
<td>14</td>
<td>2.71</td>
<td>0.71</td>
</tr>
<tr>
<td>391 Telecommunications</td>
<td>10.64</td>
<td>8</td>
<td>0</td>
<td>3</td>
<td>11</td>
<td>17.00</td>
<td>7.00</td>
</tr>
<tr>
<td>18 Safety and occupational health management</td>
<td>11.90</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>10</td>
<td>6.33</td>
<td>0.33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>256</th>
<th>83</th>
<th>106</th>
<th>445</th>
<th>3.07</th>
<th>1.28</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8.29</td>
<td>778</td>
<td>230</td>
<td>350</td>
<td>1,358</td>
<td>3.38</td>
</tr>
</tbody>
</table>

Note: Odds were calculated after adding a small constant (0.5) to all cell frequencies so that cells with zeros could be included in the calculations.

<sup>a</sup> The odds of being overgraded versus aptly graded are equal to the number of overgraded incumbents divided by the number of aptly graded incumbents.

<sup>b</sup> The odds of being undergraded versus aptly graded are equal to the number of undergraded incumbents divided by the number of aptly graded incumbents.

Source: GAO analysis of questionnaire data.

### Relationship Between Overgrading/Undergrading and Female or Minority Representation Without Controlling for GS Grade<sup>3</sup>

While we derived the study results from models fitted to table II.1, we first considered the relationship between incumbents being overgraded, aptly graded, or undergraded and female representation (see table II.2). It should be noted that the effect of the average GS grade was not statistically eliminated, or controlled for, in this comparison. For high, medium, and low female representation, overgraded and undergraded incumbents outnumbered those appropriately graded. We computed the odds as described above, and these odds indicated the extent to which overgraded or undergraded incumbents outnumbered aptly graded incumbents. We...

<sup>3</sup> All of the results reported in this section are provided for illustrative purposes only.
computed ratios, sometimes referred to as odds ratios, by dividing the odds of overgrading or undergrading in one group of occupations by the corresponding odds in another group. For example, occupations with high female representation were about 1.62 times more likely to be undergraded rather than aptly graded when compared with occupations having medium female representation (2.09 / 1.29 = 1.62). The odds ratios shown in the table indicate sizable differences in overgrading and undergrading between occupations with high female representation and those with medium representation but only slight differences between those occupations with medium representation and those with low representation.

Table II.2: Relationship Between the Observed Sample Frequencies and Female Representation Without Controlling for GS Grade

<table>
<thead>
<tr>
<th>Female representation</th>
<th>Observed frequencies</th>
<th>Odds of over versus aptly graded</th>
<th>Odds of under versus aptly graded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overgraded</td>
<td>Aptly graded</td>
<td>Undergraded</td>
</tr>
<tr>
<td>High</td>
<td>186</td>
<td>68</td>
<td>142</td>
</tr>
<tr>
<td>Medium</td>
<td>336</td>
<td>79</td>
<td>102</td>
</tr>
<tr>
<td>Low</td>
<td>256</td>
<td>83</td>
<td>106</td>
</tr>
<tr>
<td>Total</td>
<td>778</td>
<td>230</td>
<td>350</td>
</tr>
</tbody>
</table>

Note: The blank spaces in columns indicate that we did not compute odds or odds ratios.

Table II.2 indicates that all groups of occupations tended to be overgraded versus aptly or undergraded. The focus of our study was determining whether occupations with high female representation differed significantly from those with medium or low female representation. To determine whether the differences in the sample data shown in table II.2 reflected “real” differences in the population and not simply chance or sampling fluctuations, we fit the models shown in the top section of table II.3 to the data shown in table II.1.

On the basis of the likelihood ratio chi-square values and degrees of freedom associated with the different models we fit to the observed data, the preferred model was the third one in the upper and lower sections of the table. The third model in the upper section of the table indicated a pronounced tendency for incumbents of jobs with high female representation to be undergraded. The third model in the lower section of

---

4The likelihood ratio chi-square indicates the relative fit of the various models to the data in the tables.
the table indicated a pronounced tendency for the incumbents of occupations with high minority representation to be overgraded.

Table II.3: Hierarchical Models Tested to Examine Relationships Between Observed Sample Frequencies and Female and Minority Representation

<table>
<thead>
<tr>
<th>Representation</th>
<th>Model number</th>
<th>Degrees of freedom</th>
<th>Likelihood ratio chisquare</th>
<th>Probability</th>
<th>Pseudo-R² a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>F1 {F} {D}</td>
<td>4</td>
<td>37.18</td>
<td>&lt; 0.001</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>F2 {F} {D} {F,D}</td>
<td>2</td>
<td>5.63</td>
<td>0.06</td>
<td>0.849</td>
</tr>
<tr>
<td></td>
<td>F3 {F} {D} {F,D,U}</td>
<td>3</td>
<td>8.57</td>
<td>0.04</td>
<td>0.770</td>
</tr>
<tr>
<td></td>
<td>F4 {F} {D} {FD}</td>
<td>0</td>
<td>0.00</td>
<td>1.00</td>
<td>1.000</td>
</tr>
<tr>
<td>Minority</td>
<td>M1 {M} {D}</td>
<td>4</td>
<td>18.23</td>
<td>&lt; 0.001</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>M2 {M} {D} {M,D}</td>
<td>2</td>
<td>0.98</td>
<td>0.61</td>
<td>0.946</td>
</tr>
<tr>
<td></td>
<td>M3 {M} {D} {M,D,U}</td>
<td>3</td>
<td>1.57</td>
<td>0.67</td>
<td>0.914</td>
</tr>
<tr>
<td></td>
<td>M4 {M} {D} {MD}</td>
<td>0</td>
<td>0.00</td>
<td>1.00</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Legend

F = Female representation (high, medium, or low).
D = Difference between actual GS grade and questionnaire grade (overgraded, aptly, or undergraded).
M = Minority representation (high, medium, or low).

Note: The subscripts H, U, and O represent dummy variables, which contrast (1) high female or minority representation with medium and low, (2) undergraded with aptly and overgraded, and (3) overgraded with aptly and undergraded, respectively.

aPseudo-R² is calculated, following Goodman (1978), by subtracting the L² for a given model from the L² for the baseline model of independence (e.g., (37.18 - 5.63) / 37.18 = 0.849).

The result indicates the proportion of the variation in the dependent variable (the change in GS grade or the odds of overgrading or undergrading) that is accounted for by the factor or set of factors included in the model.

Table II.4 shows the results of fitting the models in the upper section of table II.3 to the data in table II.2. The odds ratios we derived from the preferred model (F3) indicated that in the population from which we drew our sample, incumbents of occupations with high female representation were twice as likely (i.e., 2.02 times as likely as shown in the table) to be undergraded compared with either aptly or overgraded, as incumbents of occupations with low or medium female representation. Again, we emphasize that the effect of the average GS grade was not statistically eliminated, or controlled for, when making this comparison.
Table II.4: Relationship Between the Expected Frequencies and Female Representation Without Controlling for GS Grade

<table>
<thead>
<tr>
<th>Model</th>
<th>Female representation</th>
<th>Expected frequencies</th>
<th>Odds of over versus aptly graded</th>
<th>Odds of under versus aptly graded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Overgraded</td>
<td>Aptly graded</td>
<td>Undergraded</td>
</tr>
<tr>
<td>F1</td>
<td>High</td>
<td>226.87</td>
<td>67.07</td>
<td>102.06</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>296.19</td>
<td>87.56</td>
<td>133.25</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>254.94</td>
<td>75.37</td>
<td>114.69</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>778.00</td>
<td>230.00</td>
<td>350.00</td>
</tr>
<tr>
<td>F2</td>
<td>High</td>
<td>186.00</td>
<td>68.00</td>
<td>142.00</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>318.15</td>
<td>87.06</td>
<td>111.78</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>273.85</td>
<td>74.94</td>
<td>96.22</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>778.00</td>
<td>230.00</td>
<td>350.00</td>
</tr>
<tr>
<td>F3</td>
<td>High</td>
<td>196.04</td>
<td>57.96</td>
<td>142.00</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>312.76</td>
<td>92.46</td>
<td>111.78</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>269.20</td>
<td>79.58</td>
<td>96.22</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>778.00</td>
<td>230.00</td>
<td>350.00</td>
</tr>
</tbody>
</table>

Note 1: The blank spaces in columns indicate that we did not compute odds or odds ratios.

Note 2: Due to rounding, expected frequencies do not always add to the total.

We also considered the relationship between incumbents being overgraded, aptly graded, or undergraded and minority representation (see table II.5). It should be noted that the effect of the average GS grade was not statistically eliminated, or controlled for, in this comparison. For high, medium, and low minority representation, the number of overgraded and undergraded incumbents outnumbered those appropriately graded. The odds shown in the table indicate the extent to which overgraded or undergraded incumbents outnumbered aptly graded incumbents. Occupations with high minority representation were 1.43 times more likely to be overgraded rather than aptly graded when compared with occupations having medium minority representation (4.48 / 3.13 = 1.43). The odds ratios shown in table II.5 indicate pronounced differences in overgrading and undergrading between occupations with high minority representation and those with medium representation but only slight differences between those occupations with medium representation and those with low representation.
### Table II.5: Relationship Between the Observed Sample Frequencies and Minority Representation Without Controlling for GS Grade

<table>
<thead>
<tr>
<th>Minority representation</th>
<th>Observed frequencies</th>
<th>Odds of over versus aptly graded</th>
<th>Odds of under versus aptly graded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overgraded</td>
<td>Aptly graded</td>
<td>Undergraded</td>
</tr>
<tr>
<td>High</td>
<td>273</td>
<td>61</td>
<td>83</td>
</tr>
<tr>
<td>Medium</td>
<td>219</td>
<td>70</td>
<td>123</td>
</tr>
<tr>
<td>Low</td>
<td>286</td>
<td>99</td>
<td>144</td>
</tr>
<tr>
<td>Total</td>
<td>778</td>
<td>230</td>
<td>350</td>
</tr>
</tbody>
</table>

Note: The blank spaces in columns indicate where we did not compute odds or odds ratios.

Table II.5 indicates that all groups of occupations tend to be overgraded versus aptly or undergraded. The focus of our study was determining whether occupations with high minority representation differed significantly from those with medium or low minority representation.

Table II.6 shows the results of fitting the models in the lower section of table II.3 to the data in table II.5. The odds ratios we derived from the expected frequencies indicate the extent of the differences that existed in the population. As previously noted, the preferred model selected was model M3, which indicated that in the population, incumbents of occupations with high minority representation were 1.64 times more likely to be overgraded rather than either aptly or undergraded. The effect of the average GS grade was not statistically eliminated, or controlled for, in this comparison.
### Table II.6: Relationship Between the Expected Frequencies and Minority Representation Without Controlling for GS Grade

<table>
<thead>
<tr>
<th>Model</th>
<th>Minority representation</th>
<th>Expected frequencies</th>
<th>Odds of over versus aptly graded</th>
<th>Odds of under versus aptly graded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Overgraded</td>
<td>Aply graded</td>
<td>Undergraded</td>
</tr>
<tr>
<td>M1</td>
<td>High</td>
<td>238.90</td>
<td>70.63</td>
<td>107.47</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>236.04</td>
<td>69.78</td>
<td>106.19</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>303.06</td>
<td>89.59</td>
<td>136.34</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>778.00</td>
<td>230.00</td>
<td>350.00</td>
</tr>
<tr>
<td>M2</td>
<td>High</td>
<td>273.00</td>
<td>61.00</td>
<td>83.00</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>221.11</td>
<td>73.99</td>
<td>116.90</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>283.89</td>
<td>95.01</td>
<td>150.10</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>778.00</td>
<td>230.00</td>
<td>350.00</td>
</tr>
<tr>
<td>M3</td>
<td>High</td>
<td>273.00</td>
<td>57.10</td>
<td>86.90</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>221.11</td>
<td>75.70</td>
<td>115.20</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>283.89</td>
<td>97.20</td>
<td>147.91</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>778.00</td>
<td>230.00</td>
<td>350.00</td>
</tr>
</tbody>
</table>

Note 1: The blank spaces in columns indicate that we did not compute odds or odds ratios.

Note 2: Due to rounding, expected frequencies do not always add to the total.

### Relationship Between Overgrading/ Undergrading Versus GS Grade and Female and Minority Representation

As depicted in figures II.1 and II.2, we determined through preliminary analyses that as the average GS grade increased (1) the odds of overgrading increased significantly and (2) the odds of undergrading decreased slightly. The trend lines through the scatterplots of points, or occupations, shown in figures II.1 and II.2 were obtained from a loglinear model, which allowed the average GS grade to be linearly related to the odds of overgrading and undergrading.\(^5\)

\(^5\)We derived the estimated odds that define the trend lines in the figures from maximum likelihood estimation procedures rather than the more common generalized least squares procedures. These estimated odds that we plotted in the two figures take into consideration the number of sampled incumbents in each occupation.
Figure II.1: Odds of Overgrading by Average GS Grade Across 37 Occupations

Note: Each plotted point represents one of the 37 occupations in our study.
Figure II.2: Odds of Undergrading by Average GS Grade Across 37 Occupations

Note: Each plotted point represents one of the 37 occupations in our study.

Table II.7 shows the loglinear models fit to the data in table II.1 to arrive at the preferred model—model 11. The first model is the baseline and asserts that the odds of overgrading and undergrading do not vary across the 37 occupations. Model 2, which allows those odds to vary linearly by average GS grade, improved significantly upon model 1 and indicates that 67 percent of the variation between the actual GS grades and grades calculated on the basis of the questionnaire was attributable to the average GS grade.\(^6\) The difference between the pseudo-R\(^2\) for model 11 and that for

\(^6\)We determine whether one model significantly improves upon another by comparing values of L\(^2\)—the likelihood ratio chi-square—which indicates the fit of the models to data or how well the expected frequencies resulting from the models correspond to the observed or sample frequencies. Since the difference in L\(^2\) values between model 1 and 2 is highly significant given the difference in degrees of freedom between the two (i.e., 698.33 – 228.10 = 470.23 with 72 – 70 = 2 degrees of freedom), we would choose model 2 as the preferred model of the two and conclude that the linear effect of the GS grade (which is present in model 2 but not in model 1) is a significant effect or one that cannot be attributed to sampling fluctuations or chance. Comparisons between other models involve similar calculations and logic.
Appendix II
Loglinear Methodology and Analysis Results

model 2 indicates that female and minority representation accounted for about 6 percent of the variation (.731 - .673 = .058). Therefore, model 11, the preferred model indicates that 73 percent of the variation between the actual GS grades and grades calculated on the basis of the questionnaire was attributable to a combination of average GS grade and female and minority representation. The remaining 27 percent of the variation was not explained by the variables in our study.

7The effects of the average GS grade and the female and minority representation operate largely independent of each other. The effect of female and minority representation would account for approximately 7 percent of the variation if introduced into the model before the GS grade.
Appendix II
Loglinear Methodology and Analysis Results

<table>
<thead>
<tr>
<th>Model number</th>
<th>Models tested</th>
<th>Degrees of freedom</th>
<th>Likelihood ratio chi-square&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Pseudo-R&lt;sup&gt;2&lt;/sup&gt; &lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[O] [D]</td>
<td>72</td>
<td>698.33</td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>[O] [G,D]</td>
<td>70</td>
<td>228.10</td>
<td>0.673</td>
</tr>
<tr>
<td>3</td>
<td>[O] [G,D] [F,D]</td>
<td>68</td>
<td>214.64</td>
<td>0.693</td>
</tr>
<tr>
<td>4</td>
<td>[O] [G,D] [M,D]</td>
<td>68</td>
<td>202.34</td>
<td>0.710</td>
</tr>
<tr>
<td>5</td>
<td>[O] [G,D] [F,D] [M,D]</td>
<td>66</td>
<td>185.26</td>
<td>0.735</td>
</tr>
<tr>
<td>6</td>
<td>[O] [G,D] [F,D] [M,D] [F,M,D]</td>
<td>64</td>
<td>183.16</td>
<td>0.738</td>
</tr>
<tr>
<td>7</td>
<td>[O] [G,D] [F,D] [M,D] [M,M,D]</td>
<td>64</td>
<td>182.77</td>
<td>0.738</td>
</tr>
<tr>
<td>8</td>
<td>[O] [G,D] [F,D] [M,D] [F,M,D]</td>
<td>64</td>
<td>181.36</td>
<td>0.740</td>
</tr>
<tr>
<td>9</td>
<td>[O] [G,D] [F,D] [M,D] [F,M,G,D]</td>
<td>64</td>
<td>181.29</td>
<td>0.740</td>
</tr>
<tr>
<td>10</td>
<td>[O] [G,D] [F,D] [M,D] [M,M,G,D]</td>
<td>64</td>
<td>180.04</td>
<td>0.742</td>
</tr>
<tr>
<td>11</td>
<td>[O] [G,D] [F,D] [M,D] [M,M,G,D]</td>
<td>68</td>
<td>188.11</td>
<td>0.731</td>
</tr>
</tbody>
</table>

Legend

O = Occupation (1 of 37 occupations covered by FES that were included in our study).
D = Difference between actual GS grade and questionnaire grade (overgraded, aptly, or undergraded).
G = Grade (GS-1 to 15).
F = Female representation (high, medium, or low).
M = Minority representation (high, medium, or low).

Subscript L indicates a linear constraint imposed upon the effect of the GS grade on overgrading or undergrading.

Subscripts H and M represent dummy variables which contrast high female or minority representation with medium and low representation, and medium female or minority representation with high and low representation, respectively.

Subscript U represents a dummy variable which contrasts undergraded with aptly and overgraded.

Subscript O indicates overgraded with aptly and undergraded.

<sup>a</sup>The likelihood ratio chi-square (L<sup>2</sup>) indicates the relative fit of the various models to the data in the table.

<sup>b</sup>Pseudo-R<sup>2</sup> is calculated, following Goodman (1978), by subtracting the L<sup>2</sup> for a given model from the L<sup>2</sup> for the baseline model of independence (e.g., (698.33 - 228.10) / 698.33 = 0.673). The result indicates the proportion of the variation in the dependent variable (the change in GS grade or the odds of overgrading or undergrading) that is accounted for by the factor or set of factors included in the model.

Table II.8 shows the expected frequencies under model 11 and the odds of overgrading and undergrading derived from them. Across the 37 occupations included in our study, of the total 1,358 positions in our sample, about 780 or 57 percent appeared to be overgraded, 352 or
26 percent appeared to be undergraded, and 232 or 17 percent appeared to be aptly graded.8

We recognize that, on the surface, these summary results could raise questions about the overall accuracy of the FES classification system. We believe that our results should be viewed with caution in this respect because we did not design our study to assess the overall accuracy of the classification system. Rather, our use of nontraditional methods, i.e., the use of the primary standard rather than occupation-specific standards on which to develop a job content questionnaire coupled with the use of the questionnaire rather than desk audits, and our sample selection methodology were designed to examine the relative assignment of grades among groups of occupations. That is, our study was designed to assess whether there were differences in the likelihood of overgrading or undergrading among groups of occupations that included higher or lower proportions of women and minorities.

We validated our design for achieving this objective; we did not validate our design for an objective of expressing an opinion on the overall accuracy of the classification system. Had we undertaken such an assessment, we would have utilized a more extensive strategy to validate the relationship between the questionnaire we developed and the results of more traditional classification tools such as desk audits, or indeed a heavier reliance directly on desk audits, which is how classification accuracy studies are usually conducted.

We also recognize that the overall extent of apparent overgrading or undergrading identified may involve some measurement error. However, we have no reason to believe that such error would be more pronounced for any particular group of occupations, for example, occupations with high female representation compared with those with medium or low representation. Thus, we do not believe that our estimates of the differences in odds of overgrading or undergrading for the groups of occupations included in our analysis have been affected by any possible measurement error. Also, the likelihood of a position being overgraded, rather than aptly graded increased as the incumbents’ GS grades increased. However, the incumbents’ grades had virtually no effect on the likelihood that a position was undergraded versus aptly graded. Accordingly, in measuring the odds of overgrading among those groups of occupations, we controlled for (statistically eliminated the effect of) the grade level

8Due to rounding, the frequency totals for overgraded, undergraded, and aptly graded positions do not add to 1,358.
effect we observed. The remaining variation in the data indicated statistically significant differences among the groups of occupations, and we report on those results.
### Table II.8: Female and Minority Representation, Average GS Grades, Expected Frequencies, and Odds of Overgraded and Undergraded Versus Aptly Graded for 37 Occupations

<table>
<thead>
<tr>
<th></th>
<th>Female Representation</th>
<th>Minority Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>High Female</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Medium Female</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Low Female</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>High Minority</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Medium Minority</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Low Minority</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

**Subtotal for occupations with high female representation**

|                  |            |         |
|                  | X          | X       |
|                  | X          |         |
|                  | X          |         |
|                  | X          | X       |
|                  | X          |         |
|                  | X          |         |
|                  | X          | X       |

**Subtotal for occupations with medium female representation**

|                  |            |         |
|                  | X          | X       |
|                  | X          |         |
|                  | X          |         |
|                  | X          | X       |
|                  | X          |         |
|                  | X          |         |
|                  | X          | X       |
## Appendix II
### Loglinear Methodology and Analysis Results

<table>
<thead>
<tr>
<th>GS occupational series</th>
<th>Average GS grade in sample</th>
<th>Expected frequencies</th>
<th>Odds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Over graded</td>
<td>Aptly graded</td>
</tr>
<tr>
<td>203 Personnel clerical and assistance</td>
<td>6.81</td>
<td>33.93</td>
<td>11.04</td>
</tr>
<tr>
<td>679 Medical clerk</td>
<td>4.45</td>
<td>4.64</td>
<td>4.88</td>
</tr>
<tr>
<td>681 Dental assistant</td>
<td>5.12</td>
<td>4.75</td>
<td>3.57</td>
</tr>
<tr>
<td>560 Budget analysis</td>
<td>11.98</td>
<td>75.78</td>
<td>4.13</td>
</tr>
<tr>
<td>318 Secretary</td>
<td>5.40</td>
<td>13.94</td>
<td>19.94</td>
</tr>
<tr>
<td>610 Nurse</td>
<td>11.20</td>
<td>16.67</td>
<td>1.34</td>
</tr>
<tr>
<td>950 Paralegal specialist</td>
<td>11.56</td>
<td>13.78</td>
<td>0.92</td>
</tr>
<tr>
<td>525 Accounting technician</td>
<td>5.62</td>
<td>2.46</td>
<td>3.16</td>
</tr>
<tr>
<td>682 Dental hygiene</td>
<td>6.56</td>
<td>16.46</td>
<td>13.22</td>
</tr>
<tr>
<td>631 Occupational therapist</td>
<td>10.50</td>
<td>9.31</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>191.72</td>
<td>63.26</td>
</tr>
<tr>
<td>360 Equal opportunity compliance</td>
<td>11.86</td>
<td>46.84</td>
<td>1.24</td>
</tr>
<tr>
<td>260 Equal employment opportunity</td>
<td>12.26</td>
<td>40.53</td>
<td>0.88</td>
</tr>
<tr>
<td>332 Computer operations</td>
<td>7.93</td>
<td>29.40</td>
<td>5.49</td>
</tr>
<tr>
<td>621 Nursing assistant</td>
<td>4.90</td>
<td>9.61</td>
<td>8.08</td>
</tr>
<tr>
<td>305 Mail and file</td>
<td>4.20</td>
<td>2.56</td>
<td>3.05</td>
</tr>
<tr>
<td>647 Diagnostic radiologic technologist</td>
<td>5.94</td>
<td>9.66</td>
<td>10.57</td>
</tr>
<tr>
<td>334 Computer specialist</td>
<td>9.67</td>
<td>24.66</td>
<td>4.23</td>
</tr>
<tr>
<td>1102 Contracting</td>
<td>8.57</td>
<td>17.38</td>
<td>5.15</td>
</tr>
<tr>
<td>510 Accounting</td>
<td>9.20</td>
<td>10.50</td>
<td>2.27</td>
</tr>
<tr>
<td>1173 Housing management</td>
<td>5.45</td>
<td>2.59</td>
<td>3.62</td>
</tr>
<tr>
<td>1371 Cartographic technician</td>
<td>6.04</td>
<td>29.15</td>
<td>30.33</td>
</tr>
<tr>
<td>1035 Public affairs</td>
<td>10.38</td>
<td>43.00</td>
<td>5.18</td>
</tr>
<tr>
<td>511 Auditing</td>
<td>10.74</td>
<td>38.62</td>
<td>3.89</td>
</tr>
<tr>
<td>690 Industrial hygiene</td>
<td>9.74</td>
<td>20.38</td>
<td>3.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>324.88</td>
<td>87.36</td>
</tr>
<tr>
<td>1896 Border patrol agent</td>
<td>10.75</td>
<td>84.62</td>
<td>3.89</td>
</tr>
<tr>
<td>85 Security guard</td>
<td>5.15</td>
<td>16.62</td>
<td>12.34</td>
</tr>
<tr>
<td>83 Police</td>
<td>5.34</td>
<td>12.93</td>
<td>19.04</td>
</tr>
</tbody>
</table>

(continued)
### Appendices II

#### Loglinear Methodology and Analysis Results

<table>
<thead>
<tr>
<th>Representation</th>
<th>Female</th>
<th>Minority</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

**Subtotal for occupations with low female representation**

**Total**
### Expected frequencies

<table>
<thead>
<tr>
<th>GS occupational series</th>
<th>Average GS grade in sample</th>
<th>Over graded</th>
<th>Aptly graded</th>
<th>Under graded</th>
<th>Total</th>
<th>Over versus aptly&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Under versus aptly&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>436 Plant protection and quarantine</td>
<td>9.13</td>
<td>11.07</td>
<td>2.48</td>
<td>2.60</td>
<td>16</td>
<td>4.46</td>
<td>1.05</td>
</tr>
<tr>
<td>1311 Physical science technician</td>
<td>5.55</td>
<td>19.43</td>
<td>25.79</td>
<td>34.93</td>
<td>80</td>
<td>0.75</td>
<td>1.35</td>
</tr>
<tr>
<td>830 Mechanical engineering</td>
<td>10.95</td>
<td>36.74</td>
<td>3.34</td>
<td>3.07</td>
<td>43</td>
<td>11.00</td>
<td>0.92</td>
</tr>
<tr>
<td>1910 Quality assurance</td>
<td>9.95</td>
<td>17.08</td>
<td>2.55</td>
<td>2.52</td>
<td>22</td>
<td>6.70</td>
<td>0.99</td>
</tr>
<tr>
<td>460 Forestry</td>
<td>8.25</td>
<td>11.62</td>
<td>4.03</td>
<td>4.50</td>
<td>20</td>
<td>2.88</td>
<td>1.12</td>
</tr>
<tr>
<td>1320 Chemistry</td>
<td>10.88</td>
<td>14.52</td>
<td>1.37</td>
<td>1.26</td>
<td>17</td>
<td>10.60</td>
<td>0.92</td>
</tr>
<tr>
<td>457 Soil conservation</td>
<td>8.75</td>
<td>10.34</td>
<td>2.80</td>
<td>3.02</td>
<td>16</td>
<td>3.69</td>
<td>1.08</td>
</tr>
<tr>
<td>475 Agricultural management</td>
<td>9.21</td>
<td>9.83</td>
<td>2.12</td>
<td>2.21</td>
<td>14</td>
<td>4.68</td>
<td>1.04</td>
</tr>
<tr>
<td>391 Telecommunications</td>
<td>10.64</td>
<td>9.25</td>
<td>0.98</td>
<td>0.92</td>
<td>11</td>
<td>9.44</td>
<td>0.94</td>
</tr>
<tr>
<td>18 Safety and occupational health</td>
<td>11.90</td>
<td>9.18</td>
<td>0.52</td>
<td>0.45</td>
<td>10</td>
<td>17.65</td>
<td>0.87</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>263.33</th>
<th>81.25</th>
<th>102.48</th>
<th>445</th>
<th></th>
<th>3.24</th>
<th>1.26</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8.29</td>
<td>779.93</td>
<td>231.87</td>
<td>351.85</td>
<td>1,358</td>
<td>3.37</td>
<td>1.52</td>
</tr>
</tbody>
</table>

<sup>a</sup>The odds of being overgraded versus aptly graded are equal to the number of overgraded incumbents divided by the number of aptly graded incumbents.

<sup>b</sup>The odds of being undergraded versus aptly graded are equal to the number of undergraded incumbents divided by the number of aptly graded incumbents.

<sup>c</sup>Due to rounding, totals for overgraded, aptly graded, and undergraded do not add to 1,358.

Table II.9 provides coefficients that indicate the direction and magnitude of the different effects included in model 11. The table shows that the relationship between the average GS grades of occupations and high minority representation and the odds of incumbents being overgraded was statistically significant, while only high female representation was related in a significant way to the odds of incumbents being undergraded.

The odds ratio 1.64 tells us that as the average GS grade increased across occupations by one grade, the odds of incumbents being overgraded increased by a factor of 1.64; that is, the likelihood of being overgraded in an occupation in which the average grade of incumbents was GS-10 was
1.64 times as great as the likelihood of being overgraded in an occupation in which the average grade was GS-9. Independent of that, occupations with high minority representation were 2.18 times more likely to be overgraded than occupations with low or medium minority representation. Finally, occupations with high female representation were 1.77 times more likely to be undergraded as occupations with low or medium female representation. The coefficient for the GS grade effect on undergrading was not significantly different from 1.0 (which indicates no effect), and the z-value associated with it implies that the effect can reasonably be due to chance.9

<table>
<thead>
<tr>
<th>Factor</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>Z-value</th>
<th>Odds ratio</th>
<th>Lower 95 percent confidence interval estimate</th>
<th>Upper 95 percent confidence interval estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Effect on overgrading</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS grade</td>
<td>0.497</td>
<td>0.038</td>
<td>13.24</td>
<td>1.64</td>
<td>1.53</td>
<td>1.77</td>
</tr>
<tr>
<td>High minority representation</td>
<td>0.780</td>
<td>0.153</td>
<td>5.10</td>
<td>2.18</td>
<td>1.62</td>
<td>2.94</td>
</tr>
<tr>
<td><strong>Effect on undergrading</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS grade</td>
<td>-0.071</td>
<td>0.043</td>
<td>-1.65</td>
<td>0.93</td>
<td>0.86</td>
<td>1.01</td>
</tr>
<tr>
<td>High female representation</td>
<td>0.570</td>
<td>0.145</td>
<td>3.94</td>
<td>1.77</td>
<td>1.33</td>
<td>2.35</td>
</tr>
</tbody>
</table>

Note: Coefficients, standard errors, and z-values are taken from the Loglinear Program in SPSS.9

9Coefficients are logged coefficients and indicate how much the natural logarithm of the odds of overgrading or undergrading differ across GS grades or female and minority representation.

9Odds ratios are derived directly from coefficients by taking antilogarithms. They are more readily interpretable than coefficients because they indicate how much the odds of overgrading or undergrading increase when GS grade increases or when occupations with high female or minority representation are compared with occupations having low or medium female or minority representation.

9Z-values are the ratios of coefficients to their standard errors. When z-values exceed +1.96 or -1.96, coefficients are statistically significant at the 95 percent level of confidence.
Appendix III

Comments From the Office of Personnel Management

Note: GAO comments supplementing those in the report text appear at the end of this appendix.

See pp. 7-9 and 61-62.

Ms. Nancy Kingsbury, Director
Federal Human Resource Management Issues
General Government Division
General Accounting Office
Washington, DC 20548

Dear Ms. Kingsbury:

This constitutes our comments on the General Accounting Office (GAO) report entitled Federal Job Classification: Comparison of Job Content to Grades Assigned in Selected Occupations which you provided to my staff August 25, 1995. In his letter of January 26, 1990, announcing the opening of this study, your predecessor, Mr. Bernard L. Ungar, stated that the Congress had asked the General Accounting Office "to determine if the federal job classification system is free of sex bias and meets current pay equity needs." He went on to say that GAO would "focus our work primarily on the design and development of the Factor Evaluation System." This study, conducted over the last 3 years, represents a significant effort to scrutinize the fairness of one of our classification systems.

We are pleased to see that your study finds no systemic bias in the Factor Evaluation System (FES). The GAO study concludes, however, that it has "identified differences in the grading of positions in occupations with high representations of women or minorities." At the same time, we note that the study contains numerous cautionary statements. For example:

- On page 8, the GAO study notes that it was not designed to "approximate the number of positions that appeared to be overgraded, undergraded, or aply graded for the portion of the federal workforce to which the results of [GAO's] study are generalizable;"

- On page 10, a footnote states that the questionnaire used to gather data "was not designed as and is not a valid substitute for traditional classification methods. Therefore the questionnaire should not be used to draw definitive conclusions regarding overgrading or undergrading of individual positions."

- On page 60, it states that GAO feels that its "results should be viewed with caution in this respect because [GAO] did not design [its] study to assess the overall accuracy of the classification system."

At several points over the course of the study, OPM expressed its strong disagreement with the methodology used. We continue to believe the methodology is insufficient to support the study’s findings. Our most immediate concerns include the following:
Appendix III
Comments From the Office of Personnel Management

Ms. Nancy Kingsbury

2.

*Use of the Primary Standard*

The FES Primary Standard was never intended to be used as it was in this study. It was designed to be used as a “standard for standards,” an overall outline within which more specific standards would be developed, not as a basis for evaluating individual positions. Indeed, we do not permit agencies to use it unsupported by an occupation-specific standard. This use of the Primary Standard outside of its designed purpose inevitably introduces anomalies and uncertainties into the study and is not reliable in determining proper grades. We believe this translates into both overgradings and undergradings that are purely a result of the study methodology and human nature, not agency classification error or systemic problems.

*Use of a less rigorous methodology to evaluate a more rigorous methodology*

GAO used a less rigorous job evaluation methodology than the agencies used. The agency decisions on the grades of jobs the study used were arrived at using officially reviewed position descriptions evaluated with occupationally specific FES standards; these agency classification decisions were made by personnel specialists specifically familiar with the work of the agency and the analytical requirements of the FES format. This is a very focused process that applies specific measures to specific duties.

GAO used generic questionnaires based on the Primary Standard, from which employees and supervisors, unfamiliar with FES ground rules, selected generic phrases that were not in context. Given these differences between the methods used in the study and the methods used by the agencies, we must give greater credibility to the agency-derived grades.

These limitations of the study affect the credibility of its data and the manner in which the data can be used.

*Disagreement between the study findings and other evidence*

A number of the findings of the study, such as the extent of the undergradings found in some occupations, contradict our daily experience and evidence from other sources. We are aware of various agency studies, conducted over time and using dependable methods, that show considerably lower levels of misclassification. These studies and other less formal evidence appear to be highly inconsistent with the GAO study. For that matter, OPM itself uses essentially the same methodology used by the agencies in determining legally authoritative grades in its classification appeals program.

Significantly, our appeals data indicate much lower misclassification rates than the GAO study. Over the past 10 years, we have found that 9 percent of the positions appealed to OPM are undergraded and 85 percent are correctly graded. These are a special subset of positions, and it is not possible to extrapolate Governmentwide classification accuracy from this sample. Nevertheless, we find it unlikely that employees who believe strongly they are misclassified would be misclassified less often as a group than employees in general.
Appendix III
Comments From the Office of Personnel Management

Ms. Nancy Kingsbury

In addition, some findings of the GAO study, such as unusual levels of over and undergrading within the same occupation, are left largely unexamined in the report and seem likely to be the result of anomalies in the study methodology rather than findings that would stand up under close examination.

Statistical emphasis

The study’s extensive statistical discussion gives an air of precision that is not supportable by the fundamental structure of the overall study design. There are several uninvestigated variables mentioned in the study, any one of which might account for some or all of the differences between the study results and the agency classification results. The report points out some of these (e.g., differences between the job content of agency position descriptions and the GAO questionnaire), but explanations for some are left unaddressed.

OPM plans to further explore the issue of classification accuracy

The report mentions the National Performance Review suggestion that the current classification system should be abandoned in favor of more flexible, broad banded systems. It then points out that policymakers should closely monitor any new systems to ensure that jobs are classified so that pay is equal for substantially equal work and that unintended disparities are guarded against. We note that the NPR recommendation focuses on ease of use and is not based on or directed at any findings of gender bias or what is referred to as "unintended disparities" in the report's conclusions.

Based on the NPR recommendation, the Administration is considering legislation that, among other things, will permit agencies to adopt alternative methods of position classification such as broad banding as well as continuing the current fifteen grade system. In anticipation of a Governmentwide broad banding authority, as well as in regard to current classification authority, OPM's newly designed oversight program will have a major focus on ensuring that current and new classification systems advance the merit principles of equal pay and the efficient and effective use of the federal workforce. In addition to conducting installation evaluations and leading a Governmentwide effort to improve internal human resources management accountability, this office will conduct several major policy reviews each year. For fiscal year 1996, one of our chief priorities is to conduct a classification review that would help facilitate the implementation of broad pay banding. For this reason, one of the options we are considering is the one you describe, a review of the accuracy of "border grades," which are those grades most likely to be placed at the lower and upper limits of any newly created pay bands. Our team is examining this and a number of other options for this review. We expect to decide on the classification review design by the end of fiscal year 1995 and begin work on the review in early fiscal year 1996.
Appendix III
Comments From the Office of Personnel Management

Ms. Nancy Kingsbury

OPM shares with GAO the need to ensure that the federal government’s classification systems and their applications are fair and unbiased. We appreciate the work GAO has done regarding this issue and thank you for this opportunity to comment on the report.

Sincerely,

James B. King
Director

See p. 12.
The following are GAO’s supplemental comments on the Office of Personnel Management’s letter dated September 8, 1995.

### GAO Comments

1. OPM said that some findings in our study, such as the overgrading and undergrading within the same occupation, were left largely unexplained. As noted in the text, the job content questionnaire was designed and validated to achieve our review objective relative to comparing groups of occupations, and we cannot attest to the questionnaire’s validity when used across GS grades within occupations, for specific occupations, or for individual positions.

2. OPM said that several uninvestigated variables were mentioned in our study, any one of which might account for some or all of the differences between the study results and the agency classification results. As noted in the text, we reported that 73 percent of the variation we found between actual GS grades and those we estimated on the basis of the questionnaire was attributable to the average GS grades and female or minority representation; the remaining 27 percent was not explained by the variables included in our study. Thus, we believe the variables in our study accounted for most of the variation.
Major Contributors to This Report

General Government Division, Washington, D.C.

Larry H. Endy, Assistant Director, Federal Management and Workforce Issues
Brenda J. Bridges, Evaluator-in-Charge
Gregory H. Wilmoth, Senior Social Science Analyst
Douglas M. Sloane, Senior Social Science Analyst
Don D. Allison, Senior Evaluator
Jennifer S. Cruise, Senior Evaluator


Bibliography


Bibliography


Bibliography


Bibliography


Bibliography


Bibliography


Bibliography


Hayes, R. “Gender Nontraditional or Sex Atypical or Gender Dominant or ... Research: Are We Measuring the Same Thing?” Journal of Vocational Behavior, Vol. 29 (1986), pp. 79-88.


Bibliography


Jeanneret, P.R. “Affidavit of Paul Richard Jeanneret.” Houston, TX: 1983.


Bibliography


Bibliography


Bibliography


Bibliography


Bibliography


Research on Evaluator Bias. Organizational Research and Development, Inc. Columbus, OH.


Bibliography


Bibliography


Related GAO Products

Pay Equity: Experiences of Canada and the Province of Ontario  

Pay Equity: Washington State’s Efforts to Address Comparable Worth  

Federal Pay: Comparisons With the Private Sector by Job and Locality  

State Department: Minorities and Women are Underrepresented in the Foreign Service  
(GAO/NSIAD-89-146, June 26, 1989).

Pay Equity: Status of State Activities  

Description of Selected Nonfederal Job Evaluation Systems  

Comments on Report on Comparable Worth by the United States Commission on Civil Rights  

Options for Conducting a Pay Equity Study of Federal Pay and Classification  

Distribution of Male and Female Employees in Four Federal Classification Systems  
(GAO/GGD-85-20, Nov. 27, 1984).

Description of Selected Systems for Classifying Federal Civilian Positions and Personnel  
(GAO/GGD-84-90, July 13, 1984).

Classification of Federal White-Collar Jobs Should Be Better Controlled  
Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are $2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20884-6015

or visit:

Room 1100
700 4th St. NW (corner of 4th and G Sts. NW)
U.S. General Accounting Office
Washington, DC

Orders may also be placed by calling (202) 512-6000 or by using fax number (301) 258-4066, or TDD (301) 413-0006.

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (202) 512-6000 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

For information on how to access GAO reports on the INTERNET, send an e-mail message with "info" in the body to:

info@www.gao.gov