

Volume 1, Number 1

November, 1986

*Personnel Assessment  
Monographs*

a publication of the

Assessment Council

of the

International Personnel Management Association

*ipmaac*

# APPLICATION OF PERSONNEL ASSESSMENT CONCEPTS AND METHODS IN JOB EVALUATION PROCEDURES

RICHARD D. ARVEY

JOHN A. FOSSUM

INDUSTRIAL RELATIONS CENTER  
UNIVERSITY OF MINNESOTA

## PERSONNEL ASSESSMENT MONOGRAPHS

James C. Johnson  
Editor

Barbara Showers  
Associate Editor

Personnel Assessment Monographs is published by the Assessment Council of the International Personnel Management Association (IPMAAC), 1617 Duke Street, Alexandria, Virginia 22314. Copies are provided without charge as a service to members of the Assessment Council. Members of IPMA and others may obtain copies by contacting Dr. Sandra Shoun at the above address.

The Monographs are intended to provide information on current topics in the field of applied personnel assessment. Manuscripts or proposals for manuscripts are solicited by the editors. Reviews of research, assessment methods, applications, innovative solutions to personnel assessment problems, and related subjects which are not generally available elsewhere are appropriate. Monographs are intended to be of value to practitioners of personnel assessment in such fields as employee selections, performance appraisal, program evaluation, recruitment, organizational assessment, and related fields. Manuscripts are reviewed by the editors and consulting editors, and accepted for publication based on the technical and professional soundness of the manuscript, and the extent to which conclusions and other information is relevant to applications in the field of personnel assessment. The editors also seek to balance topics covered to insure that issues of interest of all readers are addressed. Inquiries, proposals, and manuscripts may be sent to the current editor, James C. Johnson, 900 Plateau Parkway, Nashville, Tennessee 37205.

## Editor's Preface

Fifteen years ago, few selection specialists were concerned with classification and compensation functions in their organizations. That changed to some extent in the 1970s when the job analysis methods underlying position classification and pay-setting were found inadequate as a foundation for content-oriented employee selection procedures. These and related problems facing personnel directors and assessment specialists were described at an IPMA symposium organized by Al Maslow in the mid-1970s titled "Selection and Classification: Shall the Twain Ever Meet?", and in a 1978 Public Personnel Management article by Ollie Jensen pointing out "major conceptual morasses and misdirections of effort associated with job and personnel evaluation." Among the efforts of assessment specialists to overcome these problems were development of "multi-purpose" job analysis procedures, and involvement in establishing minimum qualification requirements.

More recent have been challenges to the basic foundations and methods of job evaluation. For some organizations, these challenges stem largely from demands that pay be equal for jobs of "comparable worth." In others, personnel directors confront employee groups and unions, managers, and others dissatisfied with job evaluation and pay decisions. One consequence is that personnel assessment specialists previously responsible only for personnel selection are increasingly being asked to assist in the preparation and review of new classification and compensation procedures, and in some cases even to assume management of such systems. Arvey and Fossum provide an introduction to this topic, outlining important problems which are commonly found in job evaluation procedures and the potential contributions of assessment specialists to solving problems which may exist in the job evaluation procedures used by their organizations. Their introduction to job evaluation is especially relevant for assessment specialists having backgrounds primarily in employee selection, and they include references to sources which cover aspects of the measurement issues they raise in greater depth.

Awarded a Ph.D. in Industrial-Organizational Psychology at the University of Minnesota in 1970, RICHARD ARVEY has served on the faculties of the Universities of Tennessee, Houston, and California at Berkeley, and is currently Professor and Acting Director of the Industrial Relations Center at the University of Minnesota in Minneapolis. He has authored more than fifty articles which have appeared in personnel, psychology, and management journals on job analysis, personnel selection, test validation, and job evaluation, and is author of the well-known text Fairness in Selecting Employees. Arvey serves as consultant to several major corporations and as instructor of the popular IPMA workshop on job evaluation.

JOHN FOSSUM served on the faculties of the University of Wyoming, the University of California at Los Angeles, and the University of Michigan before accepting his current appointment as Professor of Industrial Relations at the University of Minnesota. Best known for his work in the field of compensation, he possesses a Master's degree in Industrial Relations from Michigan State University, and a Ph.D. from the University of Minnesota.

## Application of Personnel Assessment Concepts and Methods in Job Evaluation Procedures

Richard D. Arvey                      John A. Fossum

Industrial Relations Center  
University of Minnesota

Selection specialists have rarely been involved with job classification and evaluation, largely because these functions have been handled in most organizations by an entirely separate staff responsible for compensation. But as human resource managers increasingly face challenges to their compensation practices, including hard questions about their job evaluation systems, whether they are providing equal pay for equal or in some situations comparable worth, the potential contributions of selection specialists have become apparent.

One consequence is that selection specialists are increasingly being asked to participate in designing, evaluating, and even managing such functions in their organizations. This change in the traditional role of assessment specialists provides both new opportunities and significant challenges. Few selection specialists have had formal training in compensation issues, the methods and purposes of job evaluation, or even the potential applications of assessment theory and techniques to the problems faced by classification and compensation specialists.

We hope, in this paper, to provide a start. Our purpose is to provide background information for selection/assessment professionals about job evaluation and compensation practices and to outline methods, procedures, and concepts which may have immediate applicability in the classification area. Our intent is to be didactic and provide illustrations of how selection specialists can contribute to this area rather than to provide an exhaustive or comprehensive review. Our paper is organized in the following fashion: (1) we will present a brief explanation of what job evaluation procedures are intended to accomplish and how they are used; (2) we will describe some of the more frequently used evaluation methods and why they have been utilized even more frequently within the past few years; (3) we will explore some general problems of reliability and validity in job evaluation practices and comment on the implications of these problems; (4) we will present specific ways in which persons trained in assessment and personnel selection can contribute to improved practice in job evaluation.

### Job Evaluation: A Background Sketch

Job evaluation methods were first devised shortly after WWI, a result of several changes in personnel practices. The first change was the application of job analysis by Frederick Taylor to job design in industrial engineering, and by industrial psychologists to the development of selection and placement techniques. The second change was the growth of large organizations, and the specialization of jobs. "Internal" labor markets developed in these large organizations, with the design of jobs specific to each organization. Because these jobs had no counterpart in other organizations - the external market - a new method, job evaluation, was created to establish wages for them.

Job evaluation procedures have been used primarily to help slot jobs into a wage structure which is presumed to relate to "market" wages - the wages paid by other employers competing for employees. The simplest approach was to rank order jobs in terms of their market wage, "worth," or other organizationally meaningful criterion. Nonmarket (internal labor market) jobs were slotted between ranks, based on judgments about job content by job evaluators who compared them with the content of the "market" jobs. More complex methods were quickly developed in the 1920's, particularly those known as "factor comparison" and "point factor" methods. In the 1930's, industry-wide variants of the point method such as those developed by the National Metal Trades Association were introduced. There are important distinctions between factor comparison and point methods which we will explain next.

### The Factor Comparison Method

The factor comparison method evaluates jobs using five compensable factors: skill, mental effort, physical effort, responsibility, and working conditions. Compensable factors are those characteristics of jobs presumed to reflect the tasks, duties, responsibilities, knowledges, skills, and abilities for which the employer is paying. The application of the factor comparison evaluation process, as it was conceived, begins by identifying certain jobs in the organization (usually nonexempt classifications as defined by the Fair Labor Standards Act) for which employers in the labor market compete for employees. These jobs are rank ordered for each compensable factor according to the level required by the job. Next, an independent evaluator or an evaluation team divides the wage currently paid each job into appropriate amounts for each of the compensable factors without knowledge of the assigned ranks. In the first step, then, the jobs are ranked by factors; in the second step, pay is appropriated across factors by job. Jobs are then ranked again, according to pay apportioned for each factor. A comparison is then made of the ranks of all jobs, within factors, given by each of the two methods (ranking, and pay apportioning). If there is agreement for a particular job, it is designated a "key job" and serves as an observation to be later used for slotting non-key jobs into the compensation structure. Once all of the factors in the key jobs have been "priced," the non-key jobs (primarily internal labor market) are compared factor by factor with key jobs to determine where they should be slotted and what wage rate for each factor should be extrapolated from their slot position. The entire process relies primarily on the judgment of evaluators, who are presumed to be familiar with the jobs in question or to have been made familiar with them through job analysis results.

### Point Methods

Point methods also define a set of compensable factors, frequently similar to those listed above for factor comparison methods. Each factor, however, is initially weighted with respect to its judged value or contribution to the organization. Each factor is used as a rating scale, divided into a number of levels or "degrees." Each degree is typically defined or "anchored" by a description enabling a job evaluator to match the relevant job content information with a particular degree on the rating scale. For each degree, within each factor, a certain number of "points" are defined. Normally, points for a given degree on a given factor are proportional to that factor's weight, simplifying the arithmetic calculations needed to arrive at a total point value for each job.

Point methods differ from factor comparison methods in several important ways. First, job evaluation manuals are usually used in applying point methods. The various behaviors, skills, or job demands associated with each degree are described in detail. Characteristics of each job being evaluated are compared to these definitions in order to rate and ultimately decide the "worth" of the job. Second, key jobs are also identified in point methods, but they are chosen because they are common among employers in the labor market, rather than because they have passed the ranking-pricing screen used in the factor comparison method. Third, jobs are not priced directly, as in the factor comparison method. The total points assigned each key job are plotted against the wage paid each job, forming a scatterdiagram. A regression line is calculated, and non-key jobs are slotted into the structure by placing them on or near the regression line, given their point values. The assumption underlying this procedure is that the job is worth what the market is paying for the compensable factors that make up the job.

Recently, questionnaire-based methods have been developed. Employees or job evaluators complete questionnaires describing relevant content of the job. The questionnaire results are then "scored" to arrive statistically at a job evaluation score for each job. These scores, in turn, are related to some criterion (usually market wages for certain jobs) to determine the worth of the jobs in the pay structure.

Our review of these procedures is necessarily brief; readers interested in obtaining more information about them may find Milkovich and Newman (1984) helpful.

### The Relationship Between Job Evaluation and Psychometric Scaling

Job evaluation methods are actually applications of psychometric scaling, familiar to assessment specialists, who are quite adept at scaling people. In job evaluation, however, jobs rather than people are scaled.

A great variety of tests and rating procedures are developed, subjected to psychometric evaluation, and used by assessment specialists to assess differences among people. The theory and technology needed to do so is well-developed. We identify relevant attributes or content domains to be assessed. We must often devise a sound means of reaching a "total score," either to rank candidates for appointment, or to describe "overall job performance." We might develop a rating scale to reach an "overall" score directly, or we might sum across dimensions in some fashion to obtain a total score. Notice that these kinds of measurement procedures have their counterpart in job evaluation. Although job evaluation procedures are an effort to measure jobs, the problems of content representation, weighting, reliability, and validity are common to both domains.

### **Use of Point Methods of Job Evaluation**

Job evaluation procedures have become increasingly common, particularly various versions of the point method. There are several possible reasons for increased preference for the point method. First, these scaling procedures are relatively explicit, permitting examination and review of the scales and the manner in which they are weighted to arrive at a total score for a job. The trend is toward systems which are more explicit and which permit assessment of how they are used

operationally, and away from relatively obscure systems in which employees (and sometimes even employers) don't understand how final scores or evaluations are achieved. Second, point systems may be easier to implement and to use. Finally, point systems generally permit some flexibility for local weighting of compensable factors, and definitions of degrees within factors.

#### Why is Job Evaluation Used?

Small organizations seldom use job evaluation techniques. Pay rates for employees are determined on an individual basis and depend upon knowledges, skills, and abilities; seniority; present performance level; marketability of job skills; and personal negotiating expertise. As organizations become larger, it becomes important to establish particular relationships in pay between jobs to influence the behavior of employees and to enhance their satisfaction. Behaviors that compensation programs are designed to address include attraction to the organization or a job within the organization, retention, and performance within a job. Satisfaction is assumed to depend to an extent on the degree to which the employee feels equitably treated. Perceptions of equity can be influenced both by comparisons of one's job and pay with others within the organization, as well as with those in other organizations.

Job evaluation is not designed to deal with performance motivation. It does have potential influence, however, on the organization's attraction and retention of employees through its influence on wage rates for particular jobs. It may influence perceptions of pay equity.

#### How is Job Evaluation Used?

Job evaluation procedures are often the foundation of both job classification and compensation systems. Jobs with similar point values are aggregated into "job classes." A hierarchy of classes, typically within broad occupations, is established, ranging from job classes which are lowest to highest in value.

Pay levels are typically established through use of an external referent, such as salary survey information about rates paid by other employers for similar jobs. Using this information, the relationship between points and wages is determined. To establish "external equity," points representing the joint point-wage relationships are plotted and subjected to regression analyses. By examining the wage-point trend line, organizations can determine how their overall pay policies compare to the general market. For example, the regression intercept may show that the organization is paying more, or less, than the market in general for employee wages. The organization might also find that its pay structure is steeper or flatter (regression slope) than its competition.

#### **Potential Psychometric Problems with Job Evaluation Measures**

One of us has recently described in some detail a variety of issues and problems with job evaluation methods (Arvey, in press).

1. Choice of Factors. Factors used to evaluate jobs are always an issue and sometimes controversial. While many factors (or scales) and their descriptions have evolved over the past 20-30 years and are still considered reflections of the skill, effort, responsibility, and working conditions defined earlier, critics

have argued that point plans sometimes don't include factors needed to "cover" a particular job or subset of jobs, or that factors are included which are essentially irrelevant for measuring the jobs under consideration. Assessment specialists will recognize this issue as analogous to the problem of criterion deficiency or contamination. The problem, in assessment terms, is whether the job evaluation point plan has appropriate content domain representativeness. There is a rub here. Both Title VII of the Civil Rights Act of 1964 and the Equal Pay Act of 1963 prohibit different wages for men and women when employees perform work requiring equal skill, effort, responsibility, and working conditions. When discrimination is alleged, the degree to which such factors are considered in the job evaluation procedure is an issue and, consequently, factors used in point methods ought not to drift too far from the statutory definitions. Yet, if a serious effort is made to develop point factors which represent the full range of jobs and the different kinds of tasks, duties, responsibilities demanded in them, the factors may not parallel the Equal Pay Act structure.

2. Factor Definitions and Anchor Points. A good deal of criticism has been levied at vague and imprecise definitions given both to factors and anchor points (degrees) defining levels of the various factors. Again, assessment specialists should recognize these arguments; they are similar to the criticisms of many performance appraisal scales and systems (Bernardin and Beatty, 1984). Selection specialists can play a significant role in forming more precise and perhaps more behaviorally-oriented definitions and anchor points. Determining appropriate scale values and scale lengths for the various factors are additional problems for which work is needed by persons trained in assessment methodology.

3. Weighting of Factors/Points. Review of traditional literature in this field (e.g., Belcher, 1974) reveals that compensation specialists have weighted factors and degrees in a somewhat arbitrary fashion. For example, a "Decision-Making" factor might be accorded four times as much weight as a "Working Conditions" factor through the proportional weighting schemes used. But just how these weights were determined is sometimes not well explicated nor buttressed by rationale. Selection specialists may help by providing advice on how to collect judgments needed for weighting schemes from various constituencies (e.g., management, unions, employees, consultants, etc.). While not all of these groups may agree on the weightings, at least they may be made explicit. Moreover, application of regression methods would permit assessment of the "actual" weights given to the different factors. Personnel selection specialists should recognize the similarity here with weighting test items or sections of tests. The actual weighting is a function of item variability (Schwab, 1980) and the intercorrelations with the other items. Statistical techniques can be used to "capture" the weights given to the different factors when regressed against market wages (e.g., Pierson, Koziara, & Johannesson, 1984). If one wishes to preserve the intended weighting system, standardization of the factors and explicit weighting methods must be used. In addition, systems which simply sum over the factors to arrive at a total score make it difficult to determine the actual weighting for the various factors.

4. Evaluation of Reliability. One of the most frequently encountered arguments is that job evaluation systems are unreliable because of the subjective nature of the evaluation processes. In the interim report published by the National Academy of Science, Treiman and Hartmann (1979) observed that the evidence for reliability in the job evaluation process is "not particularly encouraging." Selection specialists can provide a great deal of help in

estimating the reliability of specific job evaluation methods. Several estimates would be useful.

(1) One could assess test-retest reliability by asking raters to rate the same set of jobs on successive occasions. Cain and Green (1983) presented the results of a study of the test-retest reliabilities for Dictionary of Occupational Titles (DOT) ratings of worker functions, training times, physical demand, and working conditions for 48 occupations. The reliabilities were generally in the .80's, but there were still some considerably low estimates obtained (e.g., .04, .60, etc.).

(2) Determine estimates of inter-rater reliability. Doverspike, Carlisi, Barrett & Alexander (1983), and Doverspike and Barrett (1984), for example, calculated intraclass correlation coefficients using generalizability methods to estimate inter-rater reliability (Cronbach, Gleser, Nanda & Rajaratnam, 1972). One should be alert to the possibility of differences between reliability estimates for male and female dominated jobs. Traditional analyses have combined these jobs when calculating estimates of reliability; yet, reliability estimates could be quite different for male- and female-dominated jobs as illustrated by Arvey (in press).

Formulating research designs which permit the assessment of reliability of job evaluation methods is an important need in most organizations. It is our hunch that there may be even higher reliabilities for job evaluation methods compared to those found using psychological tests and performance measure in part because of greater heterogeneity among the jobs than found between people (and within jobs). Job evaluation systems deal with a wide range of jobs. Thus, even though there may be disagreement among raters for specific jobs (within-job variability), this effect is relatively small compared to difference across jobs. Since job evaluation includes both between and within job effects, an analysis of variance design could be quite appropriate to estimate the proportion of variance in ratings attributable to different effects.

5. Evaluation of Validity. A particularly serious issue has to do with the evaluation of the validity of alternative job evaluation methods. Two questions must be addressed. The broader question is: How well does the job evaluation methods measure the "worth" of different jobs? The more narrow question concerns the accuracy of the job evaluation system. You may recognize the similarity of these questions to those common in personnel assessment. It is particularly helpful to think of the "worth" of a job as a theoretical construct, not unlike the "Ultimate Criterion" described by Thorndike (1949) when discussing the overall worth of an employee to an organization. Once we recognize that job evaluation methods are imperfect efforts to measure some overall construct, we can begin to piece together research strategies which confirm or disconfirm our hypothesis that the measuring device has validity; i.e., that it is measuring what we intend to measure. We can bring to bear all the statistical and psychometric methods used to determine and confirm the construct validity of tests (Campbell, 1976; Arvey & Shingledecker, 1982) into the job evaluation arena. Thus, there is no one way to "validate" a job evaluation instrument. Instead, one can gather evidence from a number of sources and strategies to validate inferences concerning job worth. Several methods can be suggested:

(1) Apply different job evaluation methods to the same set of jobs. The convergent and discriminant validities of the different methods could be examined (Campbell & Fisk, 1959).

(2) Employ factor analytic methods to confirm or disconfirm the factors chosen for inclusion in the job evaluation plan. Note that a large set of jobs is needed for such a study.

(3) Correlate job evaluation ratings with employee perceptions of acceptability and correctness.

(4) Correlate job evaluation ratings with existing wages. This is the most well-known method of establishing validity. However, the wage criterion is often criticized because of its potential bias against female-dominated jobs and female wage earners. Arvey (in press) has described a number of questions about wage-job evaluation score relationships. For example, does the job evaluation instrument produce scores bearing similar relationships for both male- and female-dominated jobs? A job evaluation procedure may be more precise in its relationship to wages for male jobs than for female jobs. Similarly, does the standard error of estimate differ for male and female jobs? Are the regression lines different? The finding of different regression lines has been said to represent pay inequities which would be remedied by comparable worth approaches. In this case, male and female jobs evaluated as equal using job evaluation measures, though not necessarily identical nor similar jobs, have different predicted wages.

It is in this area that research-oriented selection specialists may make major contributions. The appropriate computation methods are not widely known or understood; different regression lines may result solely from measurement error. (cf, Arvey, Maxwell & Abrahms, 1985).

Our major point is that by applying a number of methods and strategies, it is possible to assess the construct validity of a job evaluation procedure. We do not agree with the tenet espoused by some that it is impossible to measure job worth. Psychologists have been in the business for years of defining and measuring constructs, (e.g., intelligence). The challenges associated with measuring and understanding job worth are no greater than those presented when measuring overall job performance, verbal reasoning, spatial ability, and the like. We believe that assessment specialists are particularly well-equipped to help define and improve measurement of job worth constructs because of their background and experience in doing so to assess personnel.

6. Quality Control. Longitudinal quality control is sorely needed in most organizations. Persons involved in personnel selection and in employee development recognize that predictors and training courses initially effective often lose their validity and effectiveness over time, as jobs change. The same holds with job evaluation. Two critical issues must be dealt with. First, the job evaluation procedure measures jobs at a particular point in time using a weighting scheme intended to reflect the relative importance of the compensable factors at that time. Over time, the relative importance of the factors may change, and to the extent that the weighting does not reflect current importance, the job evaluation system suffers from deficiency and/or contamination problems. Second, wages in the labor market are influenced by both supply and demand factors. As more individuals acquire the knowledges, skills and abilities that equip them for higher paying jobs, and assume such jobs, earlier "tightness" in the labor market loosens and the relative wage paid for that job declines. If market wages are used as the "worth" criterion, then failing to adjust job evaluation scores to market experience over time will result in some occupations being over- or under-paid relative to the market.

## Conclusions

This paper has demonstrated that job evaluation procedures can be substantially improved by considering and implementing methods developed and applied in personnel assessment. Perhaps the major reason that they have not previously been incorporated into job evaluation is the departmental organization of personnel functions. Staffing and compensation activities have normally been separated. While compensation specialists have been concerned about determining appropriate pay differentials between jobs, they have not been particularly concerned with the deficiency or contamination problems of their measures.

Just as the practice of personnel assessment was substantially improved by the demands of civil rights laws and regulations, so is the present practice of job evaluation being influenced by the demands of advocates of comparable worth. More attention is currently being paid to validity and reliability issues and the identification and elimination of sources of potential bias. The developed techniques of assessment specialists can substantially aid compensation practitioners and prevent the "reinvention of the wheel" in implementing job evaluation systems.

## References

- Arvey, R.D. (in press). Sex bias in job evaluation procedures. Personnel Psychology.
- Arvey, R.D., Maxwell, S.E., & Abraham, L.M. (1986). Reliability artifacts in comparable worth procedures. Journal of Applied Psychology, 70, 695-705.
- Arvey, R.D. & Shingledecker, P. (1982). Research methods in personnel management. In K.M. Rowland & G.R. Ferris (Eds.) Personnel Management. Boston: Allyn and Bacon, Inc.
- Belcher, D.W. (1974). Compensation Administration. Englewood, Cliffs, N.W.: Prentice-Hall.
- Bernardin, H.J. & Beatty, R.W. (1984). Performance appraisal: Assessing human behavior at work. Boston: Kent Publishing Comp.
- Campbell, J.P. (1976). Psychometric theory, in M.D. Dunnette (Ed.), Handbook of industrial and organizational psychology. Chicago, Ill: Rand McNally.
- Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.
- Cain, P.S. & Green, B.F. (1983). Reliabilities of selected ratings available from the Dictionary of Occupational Titles. Journal of Applied Psychology, 68, 155-165.
- Cronbach, L.J. Gleser, G.C., Nada, H. & Rajaratham, N. (1972). The dependability of behavioral measures: Theory of generalizability for scores and profiles. New York: John Wiley and Sons.

Doverspike, D., Carlisi, A., Barrett, G.V. & Alexander, R. A. (1983). Generalizability analysis of a point-method job evaluation instrument. Journal of Applied Psychology, 68, 476-483.

Doverspike, D. & Barrett, G.V. (1984). An internal bias analysis of a job evaluation instrument. Journal of Applied Psychology, 69, 648-662.

Milkovich, G.T. & Newman, J.M. (1984). Compensation. Plano, Texas: Business Publications, Inc.

Pierson, D., Koziara, K.S. & Johannesson, R. (1984). A policy-capturing application in a union setting. In Remick, H. Comparable worth and wage discrimination. Philadelphia, Penn: Temple University Press.

Schwab, D.P. (1980) Job evaluation and pay setting: Concepts and practices. In E.R. Livernash (Ed.), Comparable worth: Issues and alternatives. Washington, D.C.: Equal Employment Advisory Council.

Thorndike, R.L. (1949). Personnel selection. New York: Wiley and Sons.

Trieman, D.J. & Hartman, H. (1981). Women, work, and wages. Washington, D.C.: National Academy of Sciences.

*ipmaac*

November, 1986