DIFFERENTIAL RESPONSES ON ALTERNATELY ANCHORED JOB RATING SCALES

Jimmy L. Mitchell, Lt Col, USAF

USAF Occupational Measurement Center
Occupational Survey Branch
Lackland AFB, TX 78236

A variety of rating scales have been used with job and occupational data through the years but very seldom is a rationale given for the use of a particular scale. Likewise, there have been a number of ways in which scales have been anchored but the reasons behind the choice of a 5point scale over a 7- or 9-point scale have not typically been reported.

Viteles' job psychograph was developed in 1934; it consisted of a standard set of psychological traits, each of which was to be rated by a job analyst as to its "importance" for the job being studied (Viteles; as cited in Blum & Naylor 1968; 506). The considerable influence of this pioneering work survives today in the form of trait ratings, such as are used in the Department of Labor job analysis system (Department of Labor 1972) and in the wide-spread use of 5-point importance scales (cf. Baehr 1967; McCormick, Jeanneret, & Mecham 1972).

In some of the more recently developed job analysis systems, longer scales have been used. Hemphill (1959) in his study of executive positions, used a 7-point Part-of-the-Position scale with three verbal anchors. The Air Force occupational analysis program used first a 7-point scale and later a 9-point scale measuring relative time spent, with verbal anchors for each scale point (Morsh 1964, Driskill 1975). Other job analysis systems have used scales which vary in length from item to item (Scott 1963; Fine and Wiley 1971).

The literature on scaling provides few clues as to the optimum number of levels for job rating scales. However, Matell and Jacoby (1972) determined experimentally that if the number of scale levels exceeds 5, only about 60 percent of the scale will be used. They concluded that scales of no more than five to seven levels should be adequate for most measurement purposes.

Christal and Madden (1961) have raised the issue of being able to detect those jobs which would be "off scale" when compared to other jobs. This is an issue of particular interest when a large number of jobs are to be considered and one objective of measurement is to be able to distinguish between jobs which are substantially different. In such cases, a larger number of scale levels are needed to insure that the extreme jobs can be appropriately rated. Thus, in the Air Force

occupational analysis program, a 9-level scale is typically used. This gives the maximum possible discrimination in a single digit scale and provides the opportunity to detect extreme jobs in most Air Force occupational areas.

A potentially more serious problem lies in the selection of verbal anchors for the scale points of job rating scales. Christal and Madden (1961) noted that it has never been determined whether every scale point should have a verbal anchor. While most job rating scales which have been used through the years have provided such anchors, Hemphill (1959) used a 7-point scale with only three verbal anchors. Cragun and McCormick (1967) used this scale with Air Force officers in a study of the reliability of job ratings and their results suggest that it had considerable reliability and was to some degree preferred by incumbents in managerial positions to characterize their jobs. Tornow and Pinto (1976) used the same scale but they compressed it to a five point scale; they provided no rationale for their modification of the Hemphill scale nor any estimate of the effect of this modification on their final data.

I have not been able to find any definitive answer to the question of the anchoring of scale points in the job analysis literature. However, in the course of gathering and analyzing data for the development of a structured job analysis instrument, I chanced on some interesting results which bear on this issue.

The instrument being developed was the Professional and Managerial Position Questionnaire (PMPQ), an experimental structured job analysis questionnaire for the study of higher level jobs (Mitchell and McCormick 1976). This 93-item questionnaire was developed in the tradition of McCormick's Position Analysis Questionnaire (PAQ) but was aimed specifically at executive and management types of positions since earlier research with the PAQ had indicated that a separate instrument for higher-level positions might be appropriate (Harris & McCormick 1973).

In this new instrument, 9-point Part-of-the-Job and Complexity scales were used with verbal anchors for every other scale point (1, 3, 5, 7, and 9). Additionally, the Complexity ratings were further anchored with behavioral examples; these behavioral examples were scaled by obtaining independent ratings of a set of examples from panels of professional and academic industrial psychologists (Mitchell 1978). Also included in the instrument were items dealing with the personal requirements for the positions, to determine such things as educational levels required, prior experience, training, etc., and a section for other information, such as the number of people supervised, etc. For these items, there were numbers, categories, or constructs which were used to anchor every point of the scale, such as years of education, numbers of employees, etc. Thus, in the same instrument, there were both alternately anchored items (Part-of-the-Job, Complexity) and items with verbal anchors for every scale point (Number supervised, etc.).

You will note that for the alternately anchored items, the 2, 4, 6, and 8 response categories are consistently lower than are the 1, 3, 5, 7 and 9 categories. This pattern is perhaps even more visible if the data are plotted as histograms.

Figure 1 gives the distribution of responses for Item 1, which asks the degree to which an incumbent schedules his or her own work or the work of others. The verbally anchored scale points are indicated in this figure by cross hatching while the unanchored scale points are shown blank. You can see that all response categories were used but that there is a marked differential in response between the anchored and the unanchored scale points.

Figure 2 displays the distribution of responses for the second item in the PMPQ; how complex are the work scheduling activities of the position? Here, the anchored scale points have not only a verbal anchor but also have one or two behavioral examples to concretely reference the level of complexity. Again, there is a marked differential in response frequency between anchored and unanchored response categories.

Figure 3 represents data from Item 89, which asks the total number of personnel in units under the supervision or management control of the incumbent. Here all response categories are concretely anchored with an interval; for this item, 3 = 10 to 25 people. As you can see from the distribution of responses displayed in this figure, this is quite a different kind of distribution. There is no marked difference across adjacent items in the systematic way seen in Figures 1 and 2. Thus, there appear to be very major differences in the way individuals respond to anchored and unanchored rating scales.

We have not yet tested to see if these are significant differences. Hopefully, this work can be done in the next few months and we can come to a more concrete conclusion. When this is done, I expect that we will seek to publish the result as a short note in one of the journals.

For the present, this unexpected result has led me to question the results of some of the earlier research. Would the results of Hemphill's landmark study of executive positions have been the same had he used a verbal anchor for all scale points rather than just three anchors across seven response categories? Would Cragun and McCormick have come to the same conclusions if they had used a Part-of-the-Postion scale which was completely anchored? Of course, there are no ready answers to these questions. We have not yet done the research needed to clarify just what is going on in these cases nor do we yet have any idea of the impact of this differential response phenonomenon on the major findings of earlier research.

What is clear is that this is a phenomemon which must be looked into; we need to learn how this type of differential response tendency impacts on occupational data and ultimately on management decisions made with these data.

The PMPQ was used to gather data on 300 positions in 45 companies, schools, and government agencies throughout the country. The sample of jobs was quite diverse and salary levels ranted from about $690 per month for an administrative assistant to over $6800 per month for an executive vice president of a major company. About 250 cases had complete data and were useable in the various types of analysis planned for the study. An analysis of the distribution of responses by item was not included in the research plan but in the course of displaying some of the data for another purpose, it was noted that some items appeared to have non-normal distributions. This led to displaying the data in such a way that the distribution of responses by scale point was visible. Table 1 gives a partial picture of this data.

The items at the top of this table are those with alternately anchored response categories. Items at the bottom have a verbal anchor for each scale point.

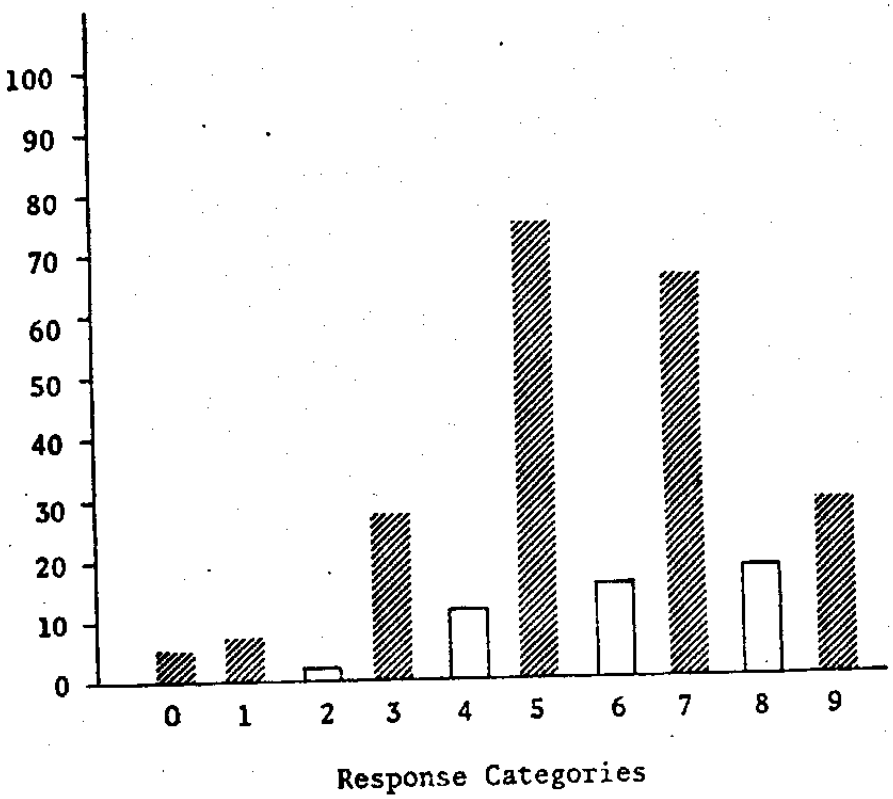

Figure 1. Distribution of responses from PMPQ Item 1. - Work Scheduling (P)

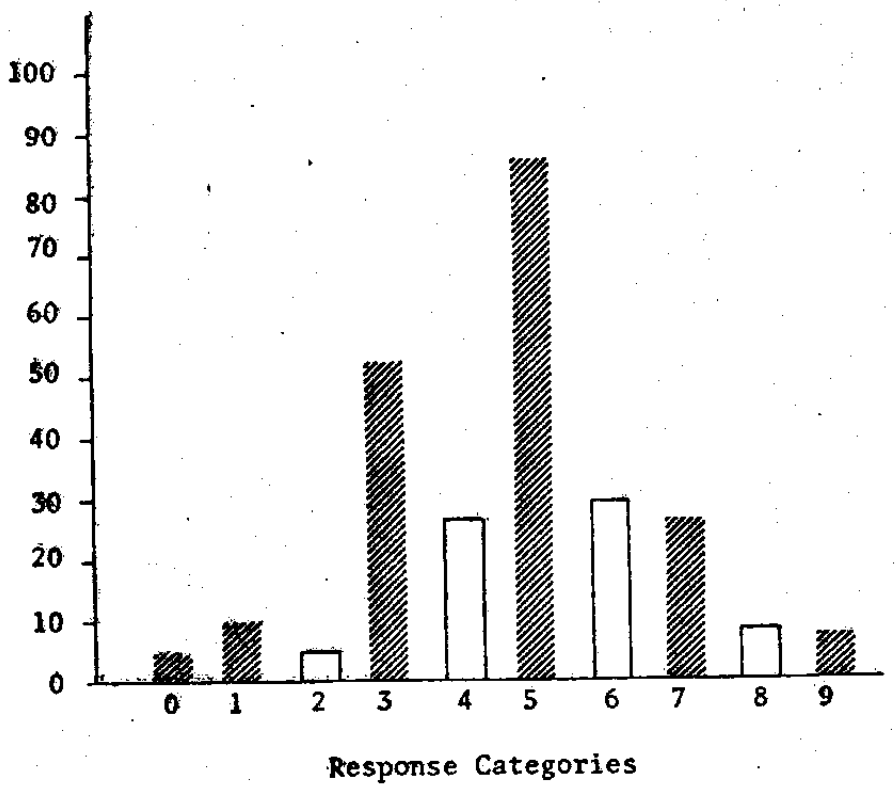

Figure 2. Distribution of responses for Item 2. - Complexity of Scheduling

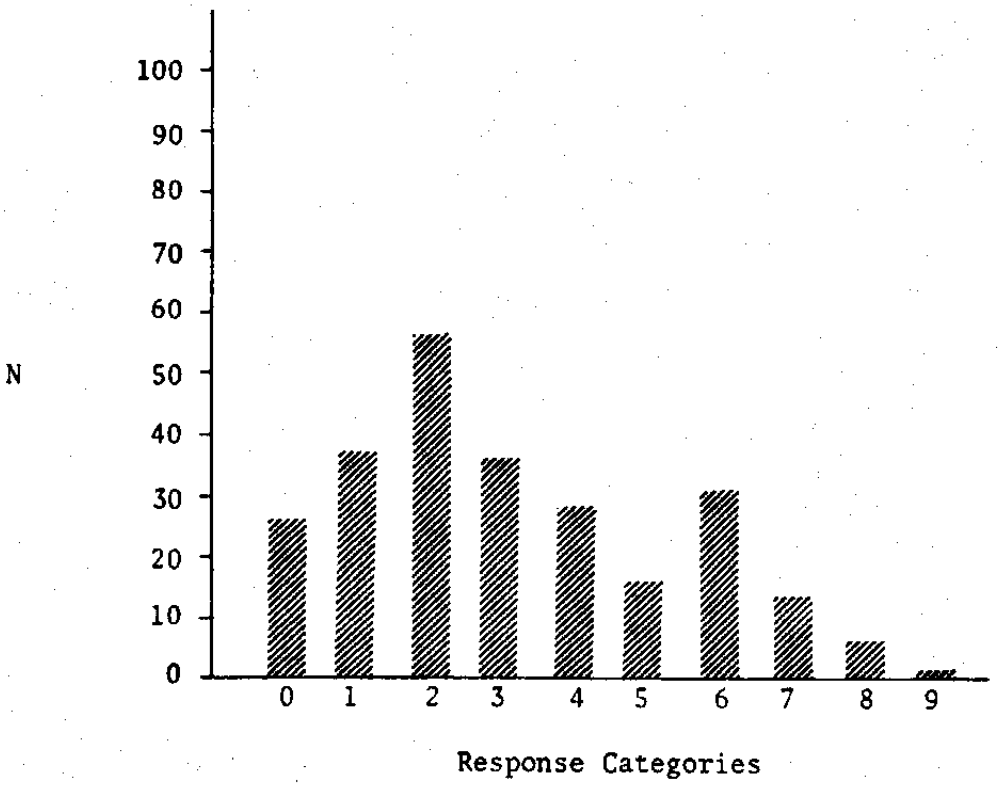

Figure 3. Distribution of responses for Item 89. - Total Number of Personnel in Units Supervised

For the present, we must assume that this type of differential response is not a desirable outcome and thus, that alternately scaled items should be avoided. Until more is known about the impact of variance in verbal anchoring such scales, scales with verbal anchors for each response category should be used. If verbal anchors cannot be developed for each scale point, then we perhaps should use a semantic differential with anchors only at the end points. It would be interesting indeed to see how our results would vary with these different anchoring systems this is an area which really could benefit from some empirical research.

BIBLIOGRAPHY

Baehr, M. E. A factorial framework for job descriptions for higher-level personnel. Industrial Relations Center, The University of Chicago, 1967.

Christal, R. E., and J. M. Madden. Air Force research on job evaluation procedures. ASD-TN-61-46. Lackland AFB, TX: Personnel Laboratory, June 1961.

Cragun, J. R., and E. J. McCormick, Job Inventory information; task and scale reliabilities and scale interrelationships. PRL-TR-67-15, Lackland AFB, TX: Personnel Research Laboratory, November 1967.

Depart of Labor. Handbook for Analyzing Jobs. Washington, D. C.: Government Printing Office, 1972.

Driskill, Walter E. Occupational Analysis in the United States Air Force. Paper presented at the Task Inventory Exchange National Symposium on Task Analysis/Inventories, The Ohio State University, Columbus, Ohio, November 18, 1975.

Fine, S. A., and W. W. Wiley. An Introduction to Functional Job Analysis. Kalamazoo, MI: The W. E. Upjohn Institute for Employment Research, 1971.

Harris, A. F. and E. J. McCormick. The analysis of rates of Naval compensation by use of a structured job analysis procedure. Report No. 3. Purdue University, Department of Psychological Sciences, September 1973.

Hemphill, J. K. Job descriptions for executives. Harvard Business Review, 1959, 37, 55-67.

Matell, M. S. and J. Jacoby. It there an optimal number of alternatives for Likert-scale items? Journal of Applied Psychology, 1972, 56 506-509.

McCormick, E. J., P. R. Jeanneret, and R. C. Mecham. A study of job characteristics and job dimensions as based on the Position Analysis Questionnaire (PAQ), Journal of Applied Psychology Monograph, 1972, 56 347-368.

Mitchell, J. L. Structured job analysis of professional and managerial positions. Unpublished doctoral dissertation, Department of Psychological Sciences, Purdue University, West Lafayette, IN, 1978.

Mitchell, J. L. and E. J. McCormick. Professional and Managerial Position Questionnaire. Department of Psychological Sciences, Purdue University, West Lafayette, IN, 1976.

Morsh, J. E., Job analysis in the United States Air Force. Personnel Psychology 1964, 17, 7-17.

Scott, W. E., Jr. The reliability and validity of a six-factor job evaluation system. Unpublished doctoral thesis, Department of Psychological Sciences, Purdue University, West Lafayette, IN, 1963.

Tornow, W. W., and P. R. Pinto. The development of a managerial job taxonomy: a system for describing, classifying, and evaluating executive positions. Journal of Applied Psychology, 1976, 61, 410-418.

Viteltes, M. S. A psychologist looks at job evaluation. Personnel, 1941, 16, 165-176, as cited in Blum and Naylor, Industrial Psychology, New York: Harper & Row 1968.