An Analysis of the Validity and Costs of Using Interviews and Questionnaires in the Study of Pay Satisfaction

Paul S. Goodman[1]

Graduate School of Industrial Administration

Carnegie-Mellon University

ABSTRACT

This paper examines the validity and costs of using a structured questionnaire and a semi-structured interview in a study of pay satisfaction. Indices from the interview measure appeared more strongly associated with the pay satisfaction measure. However, in three out of four cases both instruments contributed independently to explaining variation in pay satisfaction. The use of different coding systems were compared in terms of validity and costs.

AN ANALYSIS OF THE VALIDITY AND COSTS OF USING INTERVIEWS AND QUESTIONNAIRES IN THE STUDY OF PAY SATISFACTION

Paul S. Goodman

Graduate School of Industrial Administration

Carnegie-Mellon University

A brief review of articles concerning behavior in organizations, published in the Journal of Applied Psychology over the past five years indicates there has been little research about the methods of organizational research That is, except for validation type studies, there has not been much research about alternative strategies for conducting organizational research (see Alderfer, 1970; Alderfer and Brown, 1972, as examples of research about alternative strategies).

In addition, in more than 90 percent of the articles reviewed, the questionnaire was the dominant data collection technique. The selection of the questionnaire reflects in part the researcher's beliefs that it is more valid than other techniques (e.g., interview) and that it is less expensive to use than interviewing or observational techniques. However, few if any systematic studies in the organizational literature examine the comparative validity and the relative costs of using the questionnaire.

This paper examines the benefits and costs of using a structured questionnaire and a semi-structured interview in a study of determinants or referents affecting pay satisfaction. In studies of pay referents, the questionnaire has demonstrated acceptable validity (e.g., Andrews and Henry, 1968). The interview, on the other hand, which has not been used in this area, permits the respondent to select referents relevant to him, rather than to respond to a pre-determined set of referents. Our assumption was that people consider multiple referents in the evaluation of pay, and the combination of referents selected probably differs across respondents

This paper first compares the two instruments both in terms of their relative validity to a pay satisfaction criterion, and in their incremental

validity (Secrest, 1967). That is, the focus is not only on the comparative advantage of one instrument, but also on the effect of using both instruments on the criterion measure. If two different stimuli tap different parts of the same construct, one's ability to explain variation in the criterion measure should be enhanced.

The second issue concerns costs. There is little doubt that the total time from instrument development through coding is more expensive with an interview. However, options are available to reduce the costs of an interview and still maintain reliability and validity. The major cost in time of using interviews is the editing and coding process. In this study two methods of coding were examined -- a micro-system which coded every idea in the protocol for a given question and a global-system which coded only general categories. The trade off is simply the greater information yield under the micro-system versus the lower costs under the global-system.

Methodology

Sample

Two hundred seventeen managers from a single firm participated in this study; only three persons contacted refused to participate. The sample was selected randomly after stratifying on level and length of service.

Respondents range from those holding entry level managerial jobs to the chief operating officer. The average education for this group was 15 years. Age ranges from 25 through 65 with the average around 43 years. There are usable data on 209 respondents.

Instruments

Interview and Coding System

Four referent categories were derived from coding responses to an

open ended question in the interview which read: "You said you were (satisfied-dissatisfied) with your pay. How did you decide you were (satisfied-dissatisfied) with your pay?" After the participant provided an initial response to this item the interviewer asked: "Is there anything else that comes to your mind when you think about your pay?" The only other types of probes asked either for clarification or elaboration of a respondent's comments. The interviewers were thoroughly trained over a six week period. Analysis of the number of probes used by the four interviewers showed no significant differences across interviewers.

A coding system was developed from a sample of interviews. Two coders thoroughly trained to use this system completed the coding task, the author of this paper resolving differences. In the micro-coding system each mention of a new referent (e.g., other at my job level) was coded. The number of coded referents per interview ranged from 1 to 9, with an average of 3.3 referents per respondent. In the global-system the coders were instructed to read the entire protocol and select the two or three most salient referents. This coding was undertaken several months after the micro-system was completed.

Examples of the responses to the major item and respective codes are:

Referent Category	Operational Example of a Code
Other-Inside	"Well, we only got a 5 percent raise....but the union got 8 percent and they have much less responsibility." (inequitable)
Other-Outside	"I get about the same as I would working in a similar job outside." (equitable)
Self-Family	"I am able to provide for my family -- we live well." (equitable)

Self-Internal	"Given my length of service, and education I feel I am paid well for what I do." (equitable)

A weighting system was introduced to indicate the direction of the referent (e.g., inequitable). An equitable comparison for the elemental codes received a weight of 5, an underpayment condition 1, and an overpayment condition 4. Although overpayment and underpayment were inequitable conditions, the underpayment threshold was lower (Adams 1965) and was more likely to contribute directly to feelings of pay dissatisfaction.

Since the frequency of selecting any code was low, the general analysis procedure was to aggregate by common categories. For example, there are 15 possible codes (e.g., same level, superior, subordinates) for the Other-Inside category. If an individual selected three inside referents, his perceived inequity score would be the sum of the weights for each elemental code divided by the number of referents. Four categories are considered: Other-Inside, Other-Outside, Self-Family and Self-Internal. (See Goodman, 1974, for a complete discussion of coding categories.) The reliability of the micro-coding scheme, defined as the percent agreement of codes by category for two raters are: Other-Inside - 96%, Other-Outside - 88%, Self-Family - 96%, Self-Internal - 92%.

Questionnaire Measure

Respondents were asked to evaluate their pay (much more to much less on a 5 point scale) relevant to a series of referents. "Compared to men at my level in the company", is an item which corresponds to the Inside-Other referent. Comparison with others outside the company, friends

and neighbors, represented Outside-Other comparisons. Four items (e.g., same level, supervisor) were combined to form the Other-Inside index; three for the Outside measure, and the Self-Family, Self-Internal measures were single items.

Pay Satisfaction Index

The index is composed of four scales: (1) a Likert type item measuring satisfaction with pay (very satisfied to very dissatisfied), asked during the interview; (2) an identical pay satisfaction item included in the questionnaire which was completed approximately three to four weeks after the interview; (3) an item in the questionnaire asking how much pay there is on the job (1 minimum to 7 maximum); and (4) the difference score, found by subtracting the item asking how much pay there is from an item asking how much pay there should be (1 minimum to 7 maximum). The average intercorrelation among these scales is .62. This index has the following characteristics which aid in testing the validity of the referent categories: (1) it uses multiple items, which enhance reliability; (2) it uses different item formats and methods and the index; and (3) it uses items collected at different time periods, which should increase the independence between the referent categories and pay satisfaction.

Administration

Each respondent was first interviewed. The interview covered topics such as characteristics of his job, career plans, general life satisfaction, etc., and lasted approximately 3 hours. Approximately three weeks after the interview the same respondents received a

questionnaire which included a variety of attitudinal and personality items.

Results

Insert Table 1 about here

Table 1 presents the relationship between the interview and questionnaire pay referent measures to the pay satisfaction index. The relationships appear consistently higher with the interview measure than with the questionnaire method. The other analysis in this table concerns the incremental validity of the two measures. That is, to what extent do both instruments explain different aspects of the variation in pay satisfaction? Using a stepwise multiple regression procedure the interview entered as the most significant predictor. However, for three of the four referent categories both interview and questionnaire measures contributed to explaining variation in the criterion.

Table 2 is identical to Table 1 except that the global coding scheme was used. The relationships are essentially the same.

Insert Table 2 about here

Table 3 outlines the major activities required to transform the data from the respondents. The estimates represent total direct time allocated to the pay referent question used in this analysis. Using this table, the reader can attach his own labor costs to these figures.

Insert Table 3 about here

The interview estimates are based on the micro-coding system where the codes were developed and used for the first time. This, then, would be the most expensive option -- 5.2 hours per respondent per question. The global coding system permits reducing editing and coding by a factor of 1/3 to 1/2. Using the global system reduced time per respondent per question to approximately 3.3 hours.

When the latter coding system was used in another population to test its generalizability (Goodman 1974), the average time, using global rating, was approximately 1.5 hours per respondent. This reduction is primarily a function of the fact that the coding system was developed, and the global system was used.

Discussion

The results indicate that the interview categories are more strongly associated with the pay satisfaction criterion than the questionnaire items. Also, in three of the four categories, using both measures seemed to improve one's ability to explain variation in the criterion. These results might occur because the interview and questionnaire include different subreferents within a category such as Other-Inside. However, an effort was made to match referent categories between the two instruments. Also, some independent analyses were made of single subreferents captured by the interview (people at my level) and the corresponding item in the questionnaire. These sub-analyses indicate relationships similar to those in Table 1. Still, the matching of referent categories is not perfect and some of the results may be attributed to differences in referent categories.

Another and more plausible explanation for the results is that the two measures tap different dimensions of the same referents. For example, the interview may tend to evoke a fairly specific referent (i.e., the respondent thinks of a particular individual at his level) while the questionnaire focuses on a more general referent (e.g., people at my level). The "specific-general" dimension might account for the moderate convergent validity (correlation approximately .40 between measures) and the incremental validity for the three referent categories where both measures contributed to explaining variation in pay satisfaction.

The results also show that the interview is more expensive than the questionnaire. However, the global system substantially reduces costs without sacrificing predictive validity. Similarly, once a coding scheme is developed costs per respondent in new populations are further reduced.

The relative benefits and costs of this discussion indicate several possible guidelines. First, if the research problem requires the respondent to define dimensions in a construct space and if respondents are likely to key on different dimensions, the interview might be an appropriate instrument. Second, if time and costs permit, multiple measures might enhance our understanding of the criterion variables. Multiple measures are not only needed for assessing convergent validity but also for assessing incremental validity. Thirdly, the assessment of costs and benefits of alternative research strategies is not solely the job of large survey organizations. It is incumbent on all organizational researchers to experiment with alternative strategies and cost out the effectiveness of these strategies.

Table 1

Comparison of Interview and Questionnaire Measures with Pay Satisfaction-Micro Coding System

	Correlation Coefficient	Beta	F Values	R^2
Referent Category				
Other-Inside (n = 102)				
Interview Measure	.43*	.42	21.86*	
Questionnaire Measure	.15	.13	2.39	
				.20
Other-Outside (n = 91)				
Interview Measure	.68*	.50	43.96*	
Questionnaire	.61*	.38	25.24*	
				.58
Self-Internal (n = 61)				
Interview Measure	.53*	.45	18.19*	
Questionnaire	.38*	.32	9.37*	
				.28
Self-Family (n = 76)				
Interview Measure	.62*	.46	23.87*	
Questionnaire	.30*	.31	11.15*	
				.40

* = $p < .01$

** = $p < .05$

Table 2

Comparison of Interview and Questionnaire Measures with Pay Satisfaction-Global Coding System

	Correlation Coefficient	Beta	F. Values	R^2
Referent Category				
Other-Inside (n = 71)				
Interview Measure	.46*	.46	19.05*	.21
Questionnaire Measure	.03	.04	.18	
				.21
Other-Outside (n = 48)				
Interview Measure	.65*	.52	29.17*	
Questionnaire Measure	.61*	.45	21.91*	
				.62
Self-Internal (n = 34)				
Interview Measure	.62*	.46	13.39*	
Questionnaire Measure	.38**	.43	11.74*	
				.57
Self-Family (n = 52)				
Interview Measure	.69*	.53	.42	
Questionnaire Measure	.30**	.41	.02	
				.63

* = $p < .01$

** = $p < .05$

TABLE 3

Estimated Costs for the Development, Administration and Data Preparation for the Interview and Questions (in hours)

Research Activity	Interview	Questionnaire
Instrument development	4	1/2
Training in use of instrument	13	---
Administration of instrument	52	5
Editing and coding of instrument	1000	15
Data preparation	25	5
Hours per respondent - Micro system Coding system being developed	5.2	.10
Hours per respondent - Global system Coding system being developed	5.3	.10
Hours per respondent - Global system Coding system developed	1.5	.10

References

Alderfer, C. P. "Comparison of Questionnaire Responses With and Without Preceding Interviews." Journal of Applied Psychology, 1968, 52 335-340.

Alderfer, C. P., and Brown, D. "Designing an Emphathic Questionnaire for Organizational Research." Journal of Applied Psychology, 1972, 56, 456-460.

Andrews, I. R. and Henry, M. M. "Management Attitudes Toward Pay." Industrial Relations, 1963, 3, 29-39.

Secrest, L., "Incremental Validity" in D. Jackson and S. Messick (Eds.), Problems in Human Assessment, New York: McGraw-Hill, 1967, 368-371.

Adams, J. S. "Injustice in Social Exchange." In Berkowitz (Ed.). Advances in Experimental Social Psychology, 2, New York: Academic, 1965, 267-299.

Goodman, P. S. "An Examination of Referents Used in the Evaluation of Pay." Organizational Behavior and Human Performance. 1974 (in press).

Footnote

[1] Requests for reprints should be sent to Paul S. Goodman, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, Pennsylvania, 15213.