

# Methodological Analysis of Interview Data for Personnel Selection.<sup>1</sup>

Rhea S. DAS

Indian Statistical Institute Calcutta.

## INTRODUCTION.

The primary purpose of personnel selection is to predict, from a field of applicants, those persons who will be suitable and most successful in the particular post. Accurate prediction results in benefits to both employer and applicant. For the employer, the benefit lies in the reduction of production costs due to incompetent persons who must be fired, and due to less competent persons, who, while retained, never reach high productivity. Production costs are understood here to include not only capital expenditure, but also time and labour costs. For the

applicant, the benefit of accurate prediction lies in the fact that he will not be employed where he cannot effectively utilize his skills, and the possibility of losing the job because he is not suited for it is reduced. Inaccurate prediction, conversely, results in increased production costs for the employer and increased risk of job failure for the applicant.

Viewed in these terms, the problem now becomes one of increasing the accuracy of personnel selection methods. The interview has been the most popular and widely used method for the selection of personnel, and is often the only method used. Due to

1. Read to the Psychology and Educational Sciences Section, 44th Indian Science Congress, Calcutta, 1957.

its primary position as a selection method, it is necessary to look more critically at the interview process and the nature of the data obtained. In particular, the reliability, freedom from errors of judgement, and validity of the interview must be considered. The meaning of these three concepts in relation to the interview may be briefly mentioned. Reliability of interview judgments refers to the consistency or dependability of judgments. If the judgments are reliable, any one applicant should receive the same score or rating on two separate occasions or from two independent judges. Errors of judgment indicate that certain constant sources of bias are influencing the results. Since these may be differentially operating with respect to interviewer and candidate, they may affect the accuracy of the interview results. Examples of such errors are halo effects, errors of leniency, and the under- or over-valuing of some trait or traits in others. The validity of the interview is concerned with the accuracy of its prediction of successful personnel. In so far as the interview is valid, personnel later shown to be successful will have received favourable interview judgments.

In order to increase reliability, and to decrease errors of judgment, of the interview, quantitative interview rating forms were introduced for personnel selection purposes. The quantitative rating forms should meet this two-fold purpose because they provide a constant frame of reference for the evaluation of the candidate and a standard marking scheme. A sample form and some empirical data will be presented in this report as

a demonstration of the method and its application in personnel selection. assign marks or ratings to candidates.

#### METHOD.

The design of the quantitative interview rating form, its use in the selection interview, and a description of the data collected for the present studies will be reviewed here.

The quantitative interview rating form is designed using the following steps:

1. Analysis of qualities to be assessed by the interview in terms of the job requirements;

2. Statements of the qualities chosen so that the interview board members can readily assess them and providing a constant frame of reference for the board members; and

3. Choice of a numerical rating scale which the board members will use to

A sample interview rating form designed according to the steps given above is given in the Appendix. This form illustrates the type of forms used in the present studies.

The quantitative rating form is most appropriately used by a selection board consisting of three or more members. Each board member rates each candidate independently of the other board members. In this manner, each candidate is rated on several traits by all of the board members. The resulting data can be treated as a three-way factorial design without replication, and analyzed by analysis of variance procedures (1).

For the present studies, the data from three different personnel selection programs were analyzed. These programs will be

referred to as Programs A, B, and C: Program A utilized the interview to select a lady receptionist; Program B selected stenographers; and Program C selected laboratory technicians. The rating form, slightly modified, for Program C is given in the Appendix. The traits rated in each programs are listed

in Table 1. The selection board for each of the programs was made up of four members.

#### RESULTS.

To estimate the reliability of the interviews, the intra-class correlation (1) was used. The intraclass correlation is given by the formula.

$$r_{kk} = \frac{V_p - V_e}{V_p}$$

where  $r_{kk}$  is the reliability for k raters or interviewers,  $V_p$  is the variance for persons, and  $V_e$  the variance for error. This reliability estimate was computed for each trait on the several interview rating forms and the resulting coefficients are given in Table 1.

To determine the nature of errors of judgment in the interview, an analysis

of variance procedure was employed (1). By this procedure, variation due to interviewers, traits, and candidates can be examined for significance, and the presence of errors of judgment determined. Analysis of variance summary tables giving the sums of squares, degrees of freedom, variance estimates, F ratios, and the levels of significance, are presented for the three sets of data in table 2.

## DISCUSSION

As the most popular and widely used method in personnel selection, the interview is of importance to both employer and employee. In view of this importance, it is essential that the interview be reliable and free from errors of judgment. A methodological analysis of interview data was carried out to examine reliability and errors of judgment when quantitative rating forms were used. Use of these forms resulted in a high degree of agreement between interviewers, but did not completely control errors of judgment. These points will be discussed separately below.

Reliability of the interview ratings, estimated by the intraclass correlation, was found in general to be high in these studies. The coefficients given in Table 1 in general are higher

than those reported by other investigators. Vernon and Parry (2) report interview reliabilities of .59 and .51 for investigations utilizing summary marks only. The reliabilities reported in this paper varied for different traits. However and the ratings of some traits were relatively unreliable. Where the coefficients are low, the traits should be examined to see whether they can be feasibly evaluated during the short period of the interview. If they cannot, then they should not be included in the rating schedule. Comparison of the results of Selections A and C with those of Selection B indicates that the finer discrimination required for an eleven point rating scale does not result in greater reliability than is obtained by the five point rating scale. This factor may be considered in the design of future

interview rating schedules. While improved reliability has been found in these studies, in comparison with earlier work, the aim of investigators should be toward greater reliability. Training of interviewers, in which they are taught to look for similar characteristics and to use common criteria in evaluation, may further improve interview reliability.

Analysis of the interview ratings for errors of judgment showed that they were not entirely eliminated by the use of quantitative rating forms. However they could be detected when such forms were used. The data were treated statistically by analysis of variance for three-way factorial design without replications. This analysis revealed the presence of halo effects, leniency errors, and under- or over-valuing of traits in others. These errors may result in

ratings which the candidate does not deserve, and hence may permit unwarranted selection or rejection. Their control is necessary, therefore, to obtain the most objective estimates of the candidate on the different characteristics being rated. Halo effect refers to the tendency of a rating on one characteristic to influence ratings on other characteristics. Halo effects were shown statistically by significant interviewer-candidate interactions, and were found to be present in two of the selection programs. Errors of leniency occur because some interviewers overvalue or undervalue candidates in general, i. e., they are very generous or very strict, in their ratings. Leniency errors were evidenced by significant differences between interviewers, and were found in all the three selection

programs. The general tendency of an interviewer to overvalue or undervalue certain trait in others is another error of judgment. The under or over-valuing of traits in others is shown by significant interviewer-trait interaction, and was significant for only one of the three sets of selection data. These results suggest that the leniency errors are most likely to occur. Once such errors have been detected in the data, the ratings may be corrected by a method given by Guilford (1). Another method of controlling these errors of judgment is through training by which interviewers learn to observe such errors in their judgments and to control them.

Results of these methodological studies confirm the use of quantitative interview rating forms for

the two-fold purpose of increasing reliability and decreasing errors of judgment in the selection interview. The permit they use of a constant frame of reference by all interview board members, and at the same time allow the flexibility which many employers feel to be essential. The increased use of these forms in personnel selection would be of value in the prediction of successful personnel.

#### SUMMARY

Interview data from personnel selection programs have been subjected to a methodological analysis of reliability and errors of judgment. The results of the analysis indicate that increased reliability and statistical control of errors of judgment are possible through the use of quantitative rating forms.

#### REFERENCES

1. Guilford, J. B. *Psychometric Methods*. Second Edition. New York: McGraw-Hill, 1954.
2. Vernon, P. E. & Parry, J. B. *Personnel Selection in the British Forces*. London: University of London Press, 1949.

TABLE 1.  
Reliabilities of Interview Boards

SELECTION A: Lady Receptionist		SELECTION B: Stenographer		SELECTION C: Laboratory Technician	
Trait	Reliability	Trait	Reliability	Trait	Reliability
1 General Information	.67	1 General Information	.84	1 Technical knowledge	.84
2 Self-Expression	.84	2 Self-Expression	.80	2 Understanding Nature of Work	.95
3 Motivation	.62	3 General Attitude	.79	3 Stability in Profession	.86
4 Realistic Idea of Post	.86	4 Motivation	.63	4 Professional Promise	.93
5 Initiative	.58	5 Realistic Idea of Post	.66	5 General Suitability	.92
6 Temperament	.58	6 Initiative	.61		
7 Personal Appearance	.58	7 Temperament	.79		
8 General Suitability	.89	8 Personal Appearance	.87		

n\* = 21; k\*\* = 4  
Range of ratings:  
traits 1 to 7 = 1-4;  
trait 8 = 1-5

n\* = 15; k\*\* = 4  
Range of ratings:  
0-10

n\* = 13; k\*\* = 4  
Range of ratings:  
traits 1 to 4: 1-4  
trait 5: 1-5

\*n = number of candidates

\*\*k = number of interview board members.



**TABLE 2**  
**Summary Tables for Analysis of Variance of Interview Ratings**

Source of Variation	SECTION A			SECTION B			SECTION C			
	Sums of Squares	Degrees of Freedom	Variance of Estimate	Sums of Squares	Degrees of Freedom	Variance of Estimate	Sums of Squares	Degrees of Freedom	Variance of Estimate	
<b>1 Ignoring Individual Differences</b>										
Between interviewer visits	18.05	3	6.02	26.38	3	8.79	7.26**	5.19	3	7.73
Between interviewer	111.05	7	15.86	9.90	8	1.24	1.02	60.54	4	15.14
Interaction										
Trait Inter-	25.05	21	1.19	25.17	24	1.05	.87	4.85	12	.40
Within Sets	384.38	640	.31	697.35	504	1.21		157.02	240	.65
<b>Total</b>	<b>478.48</b>	<b>671</b>		<b>668.78</b>	<b>539</b>			<b>221.60</b>	<b>259</b>	
<b>2 Ignoring Differences Between Traits</b>										
Between viewers	18.05	3	6.02	26.38	3	8.79	16.38**	5.19	3	7.73
Between Candidates	134.42	20	6.22	12.69**	14	19.40	36.60**	705.49	12	8.79
Interaction										
Candidate	45.26	60	.75	115.89	45	2.54	5.17**	30.73	35	.86
Interaction	230.25	588	.19	256.00	480	.53		90.79	208	.44
<b>Total</b>	<b>478.48</b>	<b>671</b>		<b>668.78</b>	<b>539</b>			<b>221.60</b>	<b>259</b>	
<b>3 Ignoring Differences Between Interviewers</b>										
Between Traits	111.05	7	15.86	48.06**	8	1.24	1.72	60.54	4	15.14
Between Candidates	124.42	20	6.22	18.83**	14	19.40	26.94**	705.49	12	8.79
Interaction										
Within sets	166.90	504	.33	94.82	112	.85	1.18	47.98	195	.25
<b>Total</b>	<b>478.48</b>	<b>671</b>		<b>668.78</b>	<b>539</b>			<b>221.60</b>	<b>259</b>	

\*\*Significant at 5% level.

\*\*Significant at 1% level.

## APPENDIX

## Quantative Interview Rating Form

Initials of  
interviewer \_\_\_\_\_Name of  
Candidate \_\_\_\_\_

All candidates will be asked specific questions regarding the laboratory procedures and essential knowledge for successful work in a clinical research laboratory. When these questions have been answered, rate the quality of the answers by circling the number on the following scales which best indicates your opinion.

1. Amount of technical knowledge : is candidate well informed ?
- |   |   |   |   |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
|   |   |   |   |
- 
- |           |      |         |             |
|-----------|------|---------|-------------|
| excellent | good | average | poor        |
| very well |      |         | no          |
| informed  |      |         | information |
- 
2. Understanding nature of the work : is the candidate familiar with procedures and routine ?
- |   |   |   |   |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
|   |   |   |   |
- 
- |               |      |         |               |
|---------------|------|---------|---------------|
| excellent     | good | average | poor          |
| understanding |      |         | understanding |

Further general questions may be asked. These would include questions along the following lines : why does candidate wish to be associated with this institution ? How long does he/she expect to stay in this line of work ? What are his/her future ambitions ? Rate the quality of these answers on the following scales :

PERSONAL SELECTION

11

3. Stability in profession :    1    2    3    4  
 will candidate stay in    |    |    |    |  
 this line of work ?  


---

 most fairly unlikely definitely  
 likely likely    or    not after  
 uncertain a few years

4. Promise of candidate  
 professionally : would  
 candidate be expected    1    2    3    4  
 to advance in this type |    |    |    |  
 of work ?  


---

 definite    good    average    no  
 future    steady consistent    promise  
 promise    work    work

5. In view of the above factors, giving each the weight you  
 think important for this particular post, give a final summary  
 mark for the candidate on the following scale, by circling the  
 number which best indicates your overall opinion.

1    2    3    4    5  
 |    |    |    |    |  


---

 excellent very    good average poor  
                   good