

Validation of the Peres and Garcia Technique for Predicting Performance with Letters of Recommendation

Michael G. Aamodt, Devon A. Bryan, & Alan J. Whitcomb
Radford University

Although letters of recommendation or some form of reference checking are used by over 80% of organizations in the United States (Muchinsky, 1979), research investigating the validity of such techniques has not yielded promising results. In a study of references used in industry, Mosel and Goheen (1959) found that the validity of references was only .13. These results were supported by Browning (1968) who found the validity of references also to be .14 in predicting teaching success.

Research has identified several potential reasons for this low validity. As with the employment interview, factors other than the relevant content of the letters are used to form impressions of the applicant. For example, Cowan and Kasen (1984) found that letters referring to applicants by their first name were perceived as being more positive than letters referring to applicants by a title such as, "Mr. Jones" and Knouse (1983) found that letters of recommendation containing specific examples were evaluated more positively than letters without examples.

In an attempt to focus the attention of letter readers on the important content of the letter, Peres and Garcia (1962) developed a technique in which the traits contained in a letter of recommendation are highlighted and placed into one of five categories which were developed based on a content analysis of 625 letters of recommendation written for engineering applicants. These five categories and representative traits for each category are:

Mental Agility: Adaptable, analytical, bright, intelligent, logical, resourceful

Cooperation-Consideration: Altruistic, congenial, friendly, helpful, sincere

Dependability-Reliability: Alert, critical, dependable, methodical, prompt

Urbanity: Assured, chatty, cultured, forward, gregarious, sociable, talkative

Vigor: Active, eager, energetic, enthusiastic, independent, industrious

Unfortunately, Peres and Garcia (1962) did not attempt to validate this technique. Thus, it is the purpose of this study to investigate the reliability and validity of the technique using two separate samples.

Method

Subjects

The subjects for the first sample were 78 (39 men, 39 women) former graduate students who had complete the graduate program in psychology at Radford University. The subjects for the second samples were 26 (11 men, 15 women) graduate teaching fellows (GTFs) at Radford University. Each GTF was working toward a master's degree in psychology and had complete responsibility for teaching two sections of introductory psychology each semester.

Procedure

Graduate Student Sample. Because Radford University uses a reference rating form and actual letters of recommendation are not required, the files of over 200 former graduate students were first examined to locate those students who had at least one letter of recommendation. This process results in 78 students for whom at least one letter of recommendation as well as an overall graduate GPA were available. Two of the authors then independently:

1. Read each letter
2. Highlighted the traits in each letter
3. Used the list composed by Peres and Garcia (1962) to place each highlighted trait into one of the five categories
4. When two letters were available, the number of traits in each of the five categories was averaged across the two letters
5. To control for effects of letter length and number of traits used, the number of traits in each category was divided by the total number of traits across the five categories.

Graduate Teaching Fellow Sample. The letters of recommendation for each of the 26 GTFs were analyzed according to the procedure listed above. However, student ratings of the GTF's teaching served as the criterion rather than the GPAs used with the graduate student sample. The rating used was the final question on the rating form already used by the university that asked for an overall ratings of the instructor's performance based on a five-point rating scale with a "1" indicating poor performance and a "5" indicating excellent performance.

Results

Reliability of Letter Writers

To determine the extent to which letter writers referred to applicants with similar traits, the number of traits in each of the five categories used by each letter writer were correlated. As shown in Table 1, the coefficients across the two samples were fairly low, and in some cases, the coefficients were negative. Thus, it seems that two people writing letters for the same individual will not say the same things.

Trait Category	Sample	
	Graduate Students	Teaching Fellows
Mental agility	.12	.18
Vigor	-.03	-.58*
Dependability-Reliability	.04	.48*
Urbanity	.08	.31
Cooperation-Consideration	.15	-.17

Reliability of Letter Readers

To determine the extent to which personnel professionals reading each letter agree about the traits that are present as well as the category in which each trait belongs, the number of traits placed by the two raters in the five categories for each letter were correlated. As shown in Table 2, the coefficients were reasonably high for the first sample with the exception of the urbanity category. One of the problems encountered in the first sample was that many of the traits listed in the letters were not contained in the Peres and Garcia list. Prior to collecting data for the second sample, these new traits were added to the lists and as can be seen from Table 2, the agreement levels increased for four of the five categories.

Trait Category	Sample	
	Graduate Students	Teaching Fellows
Mental agility	.77	.91
Vigor	.86	.64
Dependability-Reliability	.70	.86
Urbanity	.53	.86
Cooperation-Consideration	.87	.96

Validity of the Trait Categories

To determine the validity of the trait categories, the percentage of traits in each of the five categories for the first sample was correlated with the student's graduate GPA and the percentage of traits in each of the five categories for the second sample was correlated with the GTF's overall student teaching ratings. As shown in Table 3, the percentage of traits in the mental agility category significantly correlated with graduate GPAs whereas the percentage of traits in the urbanity category positively correlated and the percentage of traits in the mental agility category negatively correlated with teaching ratings.

Trait Category	Criterion	
	Graduate GPA	Teaching Rating
Mental agility	.32*	-.44*
Vigor	-.08	.27
Dependability-Reliability	-.13	-.22
Urbanity	.03	.38*
Cooperation-Consideration	.04	.23

Discussion

Our findings indicate that the technique developed by Peres and Garcia (1962) shows promise as a predictor of performance. With both samples, the significant validity coefficients were more than twice the magnitude of the .13 previously found with references.

Although the Peres and Garcia categories were successful in predicting both criteria, it is possible that traits might be better classified into a different system. For example, Peres and Garcia's dependability-reliability category appears to be two separate categories; one involving dependability consisting of such items as "responsible" and "dependable" and another involving

assertiveness consisting of such traits as “tenacious,” “confident,” and “determined.” This would be an excellent topic for future research.

An interesting finding in our study was that the traits used by two letter writers to describe the same person were not highly correlated. This finding certainly makes sense if one assumes that each letter writer probably observed different aspects of the applicant’s behavior as would be the case if one of the letter writers were a professor and the other an employer.

However, even though the low agreement of letter writers is understandable, it does pose potential problems for the validity of the Peres and Garcia technique. Thus, it is important in the future to investigate such issues as the sources from which letters should be obtained as well as the optimal number of letters that should be written for each applicant. Because highlighting and categorizing traits can be a time-consuming process, future research might also want to focus on the development and validation of a trait-based checklist such as the one created by Carroll and Nash (1972).

References

- Browning, R. C. (1968). Validity of reference ratings from previous employers. *Personnel Psychology, 21*, 389-393.
- Carroll, S. J., & Nash, A. N. (1972). Effectiveness of a forced-choice reference check. *Personnel Administration, 2*, 42-46.
- Cowan, G., & Kasen, J. J. (1984). Form of reference: Sex differences in letters of recommendation. *Journal of Personality and Social Psychology, 46*, 636-645.
- Knouse, S. B. (1983). The letter of recommendation: Specificity and favorability of information. *Personnel Psychology, 36*, 331-341.
- Mosel, J. N., & Goheen, H. W. (1959). Validity of the Employment Recommendation Questionnaire II: Comparison with field investigation. *Personnel Psychology, 12*, 297-301.
- Muchinsky, P. M. (1979). The use of reference reports in personnel selection: A review and evaluation. *Journal of Occupational Psychology, 52*, 287-297.
- Peres, S. H., & Garcia, J. R. (1962). Validity and dimensions of descriptive adjectives used in reference letters. *Personnel Psychology, 15*, 79-286.