The Biographical Inventory Approach to the Selection of Teacher Trainees

S. G. Toukmanian
york university

On a utilisé un inventaire biographique de 100 questions pour prévoir le résultat de l'enseignement pratique et académique dispensé à des étudiants de première année en pédagogie. Les notes moyennes, pour l'ensemble des études et la spécialité, ainsi que les évaluations conduites par des professeurs d'étudiants en stage pédagogique, ont servi de critères pour développer trois inventaires empiriques (notant les critères) sur un échantillon de 116 stagiaires.

En utilisant un groupe séparé de 109 étudiants, on a obtenu des validités croisées significatives (de .16 à .36) entre ces clés et leurs critères respectifs. Les analyses *t*-test ont également montré que les critères-clés étaient efficaces pour la discrimination des stagiaires dont la compétence pédagogique serait ou plus ou moins élevée. De plus, on a noté que, tandis qu'aucune des variables relatives à l'admission n'était reliée au résultat dans la pratique de l'enseignement, on pouvait prévoir ce critère au moyen de deux des trois clés de l'inventaire. On a noté le besoin d'une recherche plus approfondie pour exploiter les possibilités de l'inventaire biographique comme prédicteur du résultat des futurs professeurs.

Much of the literature dealing with teacher characteristics and teaching effectiveness has provided little evidence that might be useful in the selection of candidates to teacher education programs and in the prediction of trainee future performance in the field. With respect to the latter, there is some evidence indicating that lack of success in practice teaching, i.e. low practice teaching marks, is the best available predictor of trainee performance after training (Mathis & Park, 1965; Monroe, 1952; Bedford, Note 1; Cameron, Note 2).

However, the majority of teacher training institutions admit or reject applicants on the basis of measures that have shown little or no relationship to practice teaching performance. Studies have demonstrated, for example, that the relevance of the pre-admission grade point average (Massey & Vineyard, 1958; Mathis & Park, 1965), and of scores on various tests of cognitive abilities (Shea, 1955; Carlile, 1954) is low at best; that specific trainee attributes as measured by a variety of personality and interest tests, attitude scales, etc. are of negligible predictive utility (Gowan & Gowan, 1955; Sandgren & Schmidt, 1956; Schultz & Ohlsen, 1955; Tyler, 1954); and that while pre-admission interview ratings show some relationship to marks obtained in practice teaching (Bedford, 1972) the procedure is often too demanding and impractical to implement.

As well, there is a growing recognition (Rosenshine & Furst, 1971; Ryans, 1963; Note 3) that no universally applicable criterion or criteria of students' teaching performance can ever hope to adequately account for the divergent concepts of teacher roles, the varied skills and abilities involved in teaching various grade levels, and the diversity of educational objectives held by different teacher training institutions.

The need, therefore, still exists for research leading to the development of an effective yet practical method of appraising applicants that is relevant to the purposes of a given teacher education program. Such a method has to involve an empirical assessment procedure that is capable of providing, in a relatively short period of time, large amounts of information about individuals so that a fuller characterization of differentially performing trainees can be obtained. An instrument that seems to be well qualified for these functions is the biographical inventory.

Weighted biographical information blanks have, in recent years, yielded promising results when used to predict such complex criteria as research creativity (Ellison & Taylor, 1962; Taylor, Ellison & Tucker, Note 4); job turnover in unskilled (Scott & Johnson, 1967) and clerical (Buel, 1964) employees; salary increase (Schollay, 1957); and employee theft forecasting (Rosenbaum, 1976). The procedure has also been successfully used to differentiate between administrative and research oriented scientists (Albright & Glennon, 1961), successful and less successful medical practitioners (Loughmiller, Ellison, Taylor, & Price, 1973), sales, research, and general engineers (Chaney & Owens, 1964), and between various in-service teacher criterion groups including teachers assessed as generally high and generally low on central classroom behavioral dimensions (Ryans, 1958; 1960).

The present study was conducted in an attempt to construct empirically derived biographical scoring keys for predicting the academic and practice-teaching performance of students in the first-year teacher education program at a Canadian university. A second and equally important consideration of this investigation was to determine the relative predictive efficacy of these keys as measures used to select applicants for the program.

METHOD

Subjects

The study was conducted over two academic years. From a total of 427 first-year students in education, 225 volunteered to participate in the research. Of these 116 were 1974/75 admittees. The remaining 109 were volunteers from the 1975/76 group of accepted students.

Measures

Admission Predictor Variables

Trainees were admitted to the program on the basis of (*a*) their

pre-education grade point average (GPA); (b) a pre-admission interview evaluation made by a team of interviewers from the faculty of education. Students recommended for admission were given a rating of 3, deferred students 2, and those who were refused admission were rated 1; and (*c*) referee ratings made on an evaluation form which consisted of the following 10 dimensions: demonstrated academic ability, originality and creativity, oral expressive ability, written expressive ability, ability to relate to peers, ability to relate to instructors and to others, and extent to which excited by ideas and by subject matter. Using a four-point rating scale ranging from 1 (below average or lowest 40%) to 4 (outstanding or highest 70%), two of the applicant's former academic instructors evaluated the student on each dimension. The independent ratings of the two referees were averaged to obtain a single rating per student per dimension.

The Biographical Inventory (BI)

The instrument consisted of 100 multiple-choice items. The majority of these items were selected from a biographical inventory used in an earlier study (Deirmendjian, Note 5). The items tapped a wide range of information regarding demographic data, family structure, patterns of upbringing, school experiences, personal preferences, attitudes, values, and self-evaluations.

Performance Measures

Althought no single measure of teacher effectiveness (e.g. supervisory, peer, and student evaluations; behavioral observations of teacher classroom performances; pupil gain or growth measures; etc.) has been found to be conceptually satisfactory (Getzel & Jackson, 1963), there is some indication in the literature that the most viable criterion for evaluating the practice teaching of students may be the "judgment of experts" (Bedford 1972; Odenweller, 1963; Sandiford et al., 1937). Hence, student performance in this investigation was assessed in terms of supervisory ratings obtained by use of the *Educational Assistant Evaluation Form* (EAEF) that was developed for the study. To maximize the content validity of the instrument and to avoid an arbitrary decision as to what constitutes effective teaching, the Form was developed in collaboration with the coordinators of the first-year teacher education program. The ensuing 10 bipolar seven-point scales depicted the 10 terminal objectives of the program and represented the following teacher classroom behavioral dimensions: Distant-Warm, Artificial-Genuine, Insensitive-Sensitive, Stereotyped-Resourceful, Dull-Stimulating, Uncertain-Confident, Disorganized-Systematic, Inflexible-Adaptive, Dependent–Self-directed, and Unrealistic-Realistic. A separate *Glossary*, consisting of the operational definitions of each scale, accompanied the EAEF.[1] Since the scales were highly

intercorrelated (range .55 to .77), the summed ratings across the dimensions were used to obtain a single practice-teaching score for each student.

There were also two academic performance measures: the Overall Ed-I GPA and the Major Area of Specialization GPA.

Procedure

Training

As part of their regular academic degree program all first-year students in education were required to take academic subjects in their chosen area of specialization. In addition, they were required to take a course in communicational skills, which focused on personal and interpersonal factors affecting the educational process, and a practicum which consisted of one day a week of supervised practice teaching at the elementary level. Associated with the practicum was a seminar held once a week with one or more supervising instructors or adjunct professors. The main function of the seminars was to enable students to acquire skills and information needed to solve problems arising from their classroom experiences.

Data Collection

The BI's were distributed to all first-year students shortly after they were admitted to the program. A standard set of instructions and a stamped self-addressed envelope accompanied each inventory. Students were given a brief description of the project and a statement indicating the voluntary nature of participation. They were asked to complete the inventory within two weeks and mail their answer sheet directly to the researcher.

At the close of the academic year the EAE Forms were distributed to the adjunct professors. In individually arranged meetings the 10 dimensions of the Form and the corresponding behavioral descriptions were discussed with them. They were then asked to evaluate, in collaboration with each student's host teacher, the practice-teaching performance of all the students under their supervision. The adjunct professors were blind as to who, from their group of trainees, was participating in the research; thus rater bias was minimized.

Obtaining joint adjunct-professor–host-teacher ratings had two major advantages. First, the evaluations, being based on observations of a wide variety of student activities within a particular work environment, gave a total picture of the level of student functioning. Second, since it was impractical to obtain ratings from two judges on the same set of student behaviors, the procedure yielded assessments which represented a consensus of two experts who were thoroughly familiar with the performance of the student.

Scoring Procedure

To develop an empirical scoring method for the Inventory an item analysis was carried out on the BI responses of the 116 students from the 1974/75 academic year (Sample I).[2] Biserial correlations were computed between every alternative of each item of the BI and the three criterion variables. An alternative was scored if at least 10% of the sample endorsed the alternative and if the correlation of the alternative with the criterion was .20 or greater, regardless of sign. For a correlation of .20 or greater a weight of 3 was assigned. A correlation of −.20 or greater was scored 1. Thus, a key for each criterion was established. The keys were then used to score the BI responses of the 109 students from the 1975/76 group of first-year trainees (Sample II). An estimate of the cross validities of the keys was obtained by computing Pearson product-moment correlations between the key scores and the criterion scores of Sample II.

RESULTS

The per cent agreement of each pair of referees on each of the 10 dimensions of the referee evaluation form was calculated. It was found that 93% of the paired ratings differed by 0 to 1 scale point, approximately 7% differed by 1 to 2 scale points, and only a negligible percentage (.3%) differed by 2 to 3 scale points.

The criterion intercorrelations and the validity coefficients of the BI keys and admission variables are reported in table 1.

The diagonal values in italics represent the estimated cross validities of each BI key in predicting the criterion it was empirically constructed to predict. As can be seen, all three BI keys predicted their respective criterion at the .01 and .05 levels of significance.

When the validity coefficients of the BI keys are compared with those of the admission variables, it can be seen that the most effective predictor of practice-teaching performance was the BI key constructed on the basis of the Major Area GPA criterion. The correlations between the admission variables[3] and the practice-teaching criterion ranged from −.02 to .12, indicating that these measures were essentially ineffective in predicting student teaching performance. However, in terms of the academic criteria, it was found that four of these (Pr-Ed GPA, Demonstrated Academic Ability, Oral Expressive Ability, and Written Expressive Ability) yielded validity coefficients that were higher than either the Overall or the Major Area GPA keys. As expected, the best single predictor of academic performance was the Pr-Ed GPA ($r = .59$).

To test the ability of the BI as a predictor of extremes in student performance, *t*-test analyses were also conducted on the BI scores of the top 20% and bottom 20% of Sample II. The latter group consisted of students who had received the lowest ratings in practice teaching (a total

Table 1 / Criterion Intercorrelations and Validity Coefficients of BI Keys and Admission Variables

	Criterion				
	1	2	3	*M*	*SD*
Criterion					
1. Total EAEF Score	—			54.72	10.55
2. Overall GPA Score	.28	—		6.04	.89
3. Major Area GPA Score	.30	.93	—	5.95	1.11
BI *Key Scores*					
4. Total EAEF Key Score	*.16*	.16	.17	102.98	4.05
5. Overall GPA Key Score	.14	*.36*	.27	99.58	3.52
6. Major Area GPA Key Score	.24	.30	*.22*	98.41	3.48
Admission Predictor Variables					
7. Pr-Ed GPA	.12	.59	.54	5.04	2.26
8. Demonstrated Academic Ability	.05	.52	.52	3.05	.50
9. Originality and Creativity	.02	.34	.33	3.09	.67
10. Oral Expressive Ability	.05	.39	.35	3.00	.68
11. Written Expressive Ability	.03	.45	.45	2.91	.86
12. Industry and Motivation	−.07	.25	.23	3.44	.50
13. Judgment and Maturity	−.00	.29	.25	3.48	.46
14. Leadership Ability	−.02	.33	.33	2.81	1.02
15. Ability to Relate to Peers	−.09	−.03	.00	3.17	.78
16. Ability to Relate to Instructors and others	−.04	.23	.24	3.41	.49
17. Excited by Ideas and Subject Matter	−.06	.27	.25	3.26	.52

$r_{.05} = .16$; $r_{.01} = .23$

score of 45 or less on the EAEF) and the lowest overall GPA (less than a B average) for the year. The top 20% included those with the highest academic (a B average or better) and practice- teaching evaluations (a total score of 62 or better on the EAEF). The results, reported in table 2, show that all three BI keys were effective in discriminating between the best and the poorest performing first-year trainees.

DISCUSSION

The overall results of the present study suggested that, for the particular student sample under investigation, the BI was a promising predictor of academic and practice-teaching performance and that empirically derived BI scoring keys based on these criteria could be used to discriminate between the best and the poorest functioning first-year students in education.

Table 2 / Means, Standard Deviations, and *t* Values of High and Low Functioning Trainees

BI *Score*		High Group (n = 23)	Low Group (n = 24)	t Value
EAEF Key	M	103.69	100.82	2.68**
	SD	2.86	4.24	
Overall GPA Key	M	100.69	98.65	2.00*
	SD	3.59	3.32	
Major Area GPA Key	M	99.37	96.39	3.11**
	SD	2.03	4.20	

*$p < .05$ **$p < .01$

In view of the exploratory nature of the study, the findings were encouraging in two ways. First, keeping in mind that the data were based on a selected (admitted) group of students, the results were essentially conservative estimates of the predictive power of the BI. Second, it was found that, while none of the variables considered to be important determinants of admission to the program was related to performance in practice teaching, the EAEF and the Major Area GPA keys, when validated against the same criterion, yielded validity coefficients that were significant at the .01 and .05 levels of significance.

This relative efficacy of the BI could be attributed to the multi-dimensional character of the instrument. To the extent that teaching or practice teaching implies first and foremost a human interaction that requires the application of various types and levels of interpersonal skills which are shaped by prior experiences, it seems that the BI was instrumental in producing a broad spectrum of such behavioral and experiential data. It is likely, therefore, that the biodata approach yielded several dimensions of individual differences that were related to practice teaching, and that predictors that were designed to tap specific trainee attributes or abilities (e.g. Pr-Ed GPA, oral or written expressive ability, creativity, leadership ability, etc.) failed to adequately account for the criterion dimensions.

However, the fact still remains that the validity coefficients of the BI keys accounted for only 3%–13% of the variance in the criteria. An important limiting factor in this study was the question of how one identifies or defines the effective teacher trainee in some distinctive and characteristic way. It was interesting to note, for example, that the intercorrelations between the EAEF scores and the Overall and Major Area GPA were moderate, suggesting that practice-teaching evaluations were to some extent

influenced by the student's ability to perform academically. This criterion problem was also reflected in the low validity coefficients obtained for the EAEF key. Although a great deal of thought went into the construction of the EAEF, the scale obviously needs further development. One approach would be to refine the instrument by defining its dimensions in terms of observable behaviors. Also, training the experts (adjunct professors and host teachers) in the use of the behavioral categories, and obtaining adequate reliabilities on the same judge's repeated ratings or on the ratings of several independent judges on the same trainee, could further enhance the validity of the scale and accordingly strengthen the predictive power of the BI. In short, concurrent and longitudinal research will be needed to provide improved criteria for measuring student performance in practice teaching as well as to provide a better test of the adequacy of the BI as a predictor.

However, to the degree that modest results were obtained, it is the hope of the present investigator that educational researchers will further exploit the potential of the biographical inventory approach not only as a tool for selection, but, more important, as a viable method of delineating distinctive characterizations of differentially performing teacher trainees so that a better understanding of this population can eventually be achieved.

NOTES

This study was supported by a York University Faculty of Education minor research grant. The author would like to express her gratitude to John Tibert for his assistance in the data analysis.

1. Some of the classroom behavioral dimensions and their descriptions were derived from Ryan's (1960) *Glossary of Teacher Classroom Behaviors.*

2. The scoring method was adopted from the work on biographical inventories of Dr. R. Ellison and his associates of the Institute for Behavioral Research in Creativity, Salt Lake City, Utah.

3. The pre-admission interview data were not included in the analysis. Owing to the categorical nature of the evaluations, the majority (98%) of students in the sample had received a rating of 3 on this measure.

REFERENCE NOTES

1. Bedford, C. M. *The prediction of low-practice-teaching scores at the Saskatchewan Teachers' College, Saskatoon, Saskatchewan, Canada.* Unpublished Monograph, 1959.

2. Cameron, A. A. *Teaching success of Teachers' College graduates, Saskatoon, Saskatchewan.* Unpublished study, 1957.

3. Ryans, D. G. Personal communication, 1973.

4. Taylor, C. W.; Ellison, R. L.; & Tucker, M. F. *Biographical information and the prediction of multiple criteria of success in science.* Greensboro, N.C.: Richardson Foundation, 1966.

5. Deirmendjian, S. G. *The biographical inventory: A predictor of effective interpersonal relationships among social service aides from poverty backgrounds.* Unpublished doctoral dissertation, University of Utah, 1969.

REFERENCES

Albright, L. E., & Glennon, J. R. Personal history correlates of physical scientists' career aspirations. *Journal of Applied Psychology*, 1961, *45*, 281–284.

Bedford, C. M. Screening out weak prospective teachers. *Teacher Education*, 1972, *5*, 62–73.

Buel, W. D. Voluntary female clerical turnover: The concurrent and predictive validity of a weighted application blank. *Journal of Applied Psychology*, 1964, *48*, 180–182.

Carlile, A. B. Predicting performance in the teaching profession. *Journal of Educational Research*, 1954, *47*, 641–658.

Chaney, F. B., & Owens, W. A. Life history antecedents of sales, research and general engineering interest. *Journal of Applied Psychology*, 1964, *48*, 101–105.

Ellison, R. L., & Taylor, C. W. The development and cross validation of a biographical information inventory for predicting success in science. *American Psychologist*, 1962, *17*, 391–392.

Getzels, J. W., & Jackson, P. W. The teacher's personality and characteristics. In N. E. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNally, 1963.

Gowan, J. C., & Gowan, May S. The Guilford-Zimmerman and the California Psychological Inventory in the measurement of teaching candidates. *California Journal of Educational Research*, 1955, *6*, 35–37.

Loughmiller, G. C.; Ellison, R. L.; Taylor, C. W.; & Price, P. B. Predicting career performance of physicians using the biographical inventory approach. *Journal of Vocational Behavior*, 1973, *3*, 269–278.

Massey, H. W., & Vineyard, E. E. Relationship between scholarship and first year teaching success. *Journal of Teacher Education*, 1958, *9*, 297–301.

Mathis, C., & Park, Y. H. Some factors relating to success in student teaching. *Journal of Educational Research*, 1965, *58*, 420–422.

Monroe, W. S. (Ed.), *Encyclopedia of education research*. New York: The Macmillan Co., 1952.

Odenweller, A. L. *Predicting the quality of teaching*. New York: Columbia University, Teachers College, Bureau of Publication, 1963.

Rosenbaum, R. W. Predictability of employee theft using weighted application blanks. *Journal of Applied Psychology*, 1976, *61*, 94–98.

Rosenshine, B., & Furst, N. Research on teacher performance criteria. In B. O. Smith (Ed.), *Research on teacher education: An AERA symposium*. New York: Prentice Hall, 1971.

Ryans, D. G. Some validity extension data from empirically derived predictors of teacher behavior. *Educational and Psychological Measurement*, 1958, *18*, 355–370.

Ryans, D. G. *Characteristics of teachers*. Washington, D.C.: American Council on Education, 1960.

Ryans, D. G. Teacher behavior theory and research: Implications for teacher education. *Journal of Teacher Education*, 1963, *14*, 274–293.

Sandgren, D. L., & Schmidt, L. G. Does practice teaching change attitudes toward teaching? *Journal of Educational Research*, 1956, *49*, 673–680.

Sandiford, P.; Cameron, M. A.; Conway, C. B.; & Long, J. A. *Forecasting teaching ability*. Toronto: University of Toronto Press, 1937.

Schultz, R. E., & Ohlsen, M. M. Interest patterns of best and poorest student teachers. *Journal of Educational Sociology*, 1955, *29*, 108–112.

Schollay, R. W. Personal history data as a predictor of success. *Personnel Psychology*, 1957, *10*, 23–26.

Scott, R. D., & Johnson, R. W. Use of the weighted application blank in selecting unskilled employees. *Journal of Applied Psychology*, 1967, *51*, 393–395.

Shea, J. A. The predictive value of various combinations of standardized tests and subtests for prognosis of teaching efficiency. *Catholic University of America Educational Research Monograph*, 1955, *19*, No. 5.

Tyler, F. T. The prediction of student-teaching success from personality inventories. *University of California Publications in Education*, 1954, *11*, 233–313.

Shaké G. Toukmanian is Assistant Professor of Psychology at York University, Downsview, Ont. M3J 1P3.