DOCUMENT RESUME

ED 031 742 CG 004 266

By-Prediger, Dale J.

Biographical Data Differentiating College Attenders and Nonattenders at Various Ability Levels.

American Personnel and Guidance Association, Washington, D.C.

Pub Date Mar 69

Note-17p.; Revision of a paper presented at the American Personnel and Guidance Association Convention, Las Vegas, Nevada, March 30--April 3, 1969.

EDRS Price MF-$0.25 HC-$0.95

Descriptors-Ability, *Academic Aptitude, *College Attendance, *Data Analysis, *High School Students, Males, Prediction, *Predictive Measurement, Research Projects

This study was designed to: (1) compare the contribution of various categories of biographical data (biodata) to the differentiation of college attenders and nonattenders, and (2) to assess the accuracy of biodata as a predictor of collete attendance. Subjects consisted of 20,367 high school senior males who participated in the 1960 data collection phase of Project TALENT and who responded to a 1961 follow-up questionnaire. Biodata obtained from a 394-item inventory were categorized into 14 areas. The study revealed that biodata were potent and pervasive predictors of college attendance, predictors that were not just concomitants of academic aptitude. (Author/EK)

BIOGRAPHICAL DATA DIFFERENTIATING COLLEGE ATTENDERS AND NONATTENDERS AT VARIOUS ABILITY LEVELS

Revision of paper presented at Session 386
of the 1969 APGA Convention

by
Dale J. Prediger
University of Toledo

Abstract

The major objective was to compare the contribution of various categories of biographical data (biodata) to the differentiation of college attenders and nonattenders falling within six different academic aptitude levels. The secondary objective was to determine the accuracy of predictions of college attendance based on this biodata. Subjects consisted of 20,367 high school senior males who participated in the 1960 data collection phase of Project TALENT and who responded to a 1961 follow-up questionnaire. Biodata obtained from a 394-item inventory were categorized into 14 areas. The contribution of these categories to the differentiation of attenders and nonattenders does not appear to change substantially from one ability level to another. Categories representing fruitful sources of biodata items were identified. The hit rate for attendance predictions was 76% for a cross-validation sample. This rate and the corresponding point biserial *r* of .58 were substantially higher than the results obtained when academic aptitude was used as the predictor.

BIOGRAPHICAL DATA DIFFERENTIATING COLLEGE ATTENDERS AND NONATTENDERS AT VARIOUS ABILITY LEVELS

Dale J. Prediger[1]

University of Toledo

Numerous studies have reported relationships between student biographical data (biodata) and college attendance. The types of biodata items frequently cited in the reviews by Beezer & Hjelm (1961), Berdie & Hood (1965), and Heist (1960) have included father's occupation; parents' educational level; economic status of family; cultural level of family as indicated by number of books in home, cultural activities, etc.; area of residence; size of high school; and sex of student. As would be expected, academic aptitude has also been shown to be a consistent and effective predictor of college attendance.

Many of the studies of biographical concomitants of college attendance have employed large samples and well-conceived plans for data collection. For example, Medsker & Trent (1965) obtained a carefully chosen, ad hoc sample of more than 10,000 high school graduates from 16 communities in the Midwest and Far West. Academic aptitude, high school rank, and a number of home background items were found to differentiate college attenders and nonattenders in this group. In another large scale study, Berdie & Hood (1965) have shown that a wide variety of biodata factors are also related to post-high school plans. In both of these studies, efforts were made to determine whether the relationships observed in total groups

would also be present at different academic aptitude levels. Medsker and Trent found that the differentiation of college attenders and non-attenders achieved through knowledge of father's occupation was substantial at each of three ability levels. Berdie and Hood reported relationships between college plans and sex of student, father's occupation, parents' educational level, source of family income, number of books in the home, and area of residence for Minnesota high school seniors falling in the upper 17% of the statewide norm group on academic aptitude.

Undoubtedly the most extensive studies of the relationship between biodata and college attendance have been conducted by members of the Project TALENT research staff. Numerous biodata items on the Student Information Blank (SIB) were found to differentiate college attenders and nonattenders (Flanagan, Davis, Dailey, Shaycoft, Orr, Goldberg, & Neyman, 1964). An index of socioeconomic status based on nine SIB items has also been shown to differentiate college attenders and nonattenders falling within each of four academic aptitude levels (Flanagan & Cooley, 1966).

In summary, there is considerable evidence of relationship between biodata and college attendance for unselected groups of high school students. A few studies have also shown that this relationship is present within certain ability level subgroups. However, the biodata used in these studies have generally been limited to a few indices of socioeconomic status. Whether this or any other type of biodata is consistently related to the college attendance of students falling at different ability levels remains to be determined. Information bearing on this point can carry practical implications for efforts to identify

subgroups of students for whom college attendance is or is not likely. Such information should also contribute to the understanding of the sociology of college-going.

Problem

The major objective of the study was to compare the contribution of various categories of biodata to the differentiation of college attenders and nonattenders falling within six different academic aptitude levels. The secondary objective was to assess the power of biodata as a predictor of college attendance by determining the accuracy (hit rate) of the college attendance predictions.

Method

Subjects

The records for 20,367 high school senior boys were obtained from the Project TALENT Data Bank files in the summer of 1966. These boys were all attending schools included in Project TALENT's 1960, stratified, random sample of U.S. high schools. Representativeness of the 1960 sample is discussed by Flanagan et al., (1964). An overview of the Data Bank files available in 1966 is given in a Project TALENT (1965) publication entitled "The Project TALENT Data Bank."

The sample includes only those students who chose to respond to the follow-up questionnaire mailed in the summer of 1961. Unfortunately, about 30% of the students in the 1960 sample failed to respond (Flanagan & Cooley, 1966). The absense of data on these nonrespondents may have some effect on the results summarized in this report. However, it should be noted that no attempt is being made to estimate population parameters

such as attendance rates or percentage of students responding in a given way to an item. There is no reason to believe that the nature of the relationship between biodata and college attendance would be grossly affected if data from the nonrespondents had been available.

Variables

Academic aptitude was represented by an equally weighted combination of reading comprehension and mathematics achievement test scores taken from the Project TALENT test battery. Biodata responses were obtained from the 394-item SIB. Since a number of these items covered other than biographical topics (e.g., study habits, plans, possible means of financing college, etc.), only 246 of the items were used in the analyses. For purposes of summary, these items have been organized into the 5 major categories and 14 subcategories shown in Table 1.

- - - - - - - - - - - - - - - -

Insert Table 1 about here.

- - - - - - - - - - - - - - - -

Student responses to item two of the 1961 follow-up questionnaire served to operationally define the college attendance criterion. Only those students indicating that they had entered college as full-time students were included in the college attendance group.

Design

The sample was randomly subdivided into four subsamples for purposes of data analysis. Subsample 1 (N=10,183) was used to establish academic aptitude score limits for six ability level groups of approximately equal size. The number of students and the college attendance rate for these

ability groups and for the total subsample are shown in Table 2. As would be expected, the college attendance rate differed greatly across the ability levels.

- - - - - - - - - - - - - - - -

Insert Table 2 about here.

- - - - - - - - - - - - - - - -

Item analyses were also conducted on Subsample 1 with separate analyses being performed for all students in the subsample and for students falling in each of the six ability level groups. Thus, there were a total of seven separate item analyses. In each of these, a two-by-two contingency table was formed for each SIB item response. Marked-not marked and attend-nonattend dichotomies served as the basis for these tables. Within a given ability group, an item response had to meet a chi square probability level cutoff of $p < .005$ in order for the item to be identified as differentiating attenders and nonattenders. Use of this significance level was purely a matter of convenience since, as has already been noted, the sample was not a random sample of high school senior males.

The remaining three subsamples of the total sample were used to determine the accuracy of college attendance predictions. Item responses for the SIB were weighted +1 or -1 according to direction of relationship with the criterion. Five scoring keys of different length were then formed on the basis of item analysis data obtained for all students in Subsample 1 regardless of ability level. These keys were applied to the SIB responses of Subsample 2 (N=3,387) in order to determine optimum scoring key length. The responses used in the optimum key, to be called the general key, were then scored on students in Subsample 3 (N=3,393). Equations for predicting college attendance were developed on Subsample 3 and applied to the scores

of students in Subsample 4 (N=3,404) in order to determine the accuracy of attendance predictions. The maximum likelihood classification procedures described by Cooley & Lohnes (1962) were employed.

Results

The results of the item analyses conducted on Subsample 1 are presented in Table 3. Since 2,269 item responses met the $p < .005$ cutoff criterion for the six ability levels, some method of summarization was essential. The basic units for the summary consisted of items having one or more responses meeting the cutoff criterion. Such items were separately identified for each of the

- - - - - - - - - - - - - - - -

Insert Table 3 about here.

- - - - - - - - - - - - - - - -

six ability level groups and for the total sample. Only the percentage of these items falling within the various biodata categories is shown in Table 3. Thus, for the first (lowest) ability level group, 16% of the items with responses meeting the cutoff criterion fell within the extra-curricular category. This category included 14% of the items with responses meeting the cutoff criterion in the second ability level group. Comparisons across ability levels for the other categories of biodata items can be readily made.

Perhaps the most significant aspect of the data in Table 3 is the almost complete absense of any striking trends across the six ability levels. One would expect that any effect which ability has on the relationship between biodata and college attendance would be reflected in a gradual increase or decrease across the six ability levels in the percentage of items falling within a given biodata category. The ability levels are, after all, adjacent. There is no reason to expect discontinuities in the effect of ability level on these percentages. However,

with few exceptions (the academic background category being the most notable), the fluctuations in percentages across ability levels appear to be fortuitous and indicative of little if any differential loading on the part of the biodata item categories. These results are in accordance with the results of the earlier study by Prediger (1969). In an attempt to capitalize on the uniqueness of the biodata differentiating attenders and nonattenders at a given ability level, Prediger developed separate scoring keys at each of the six levels. Use of these moderated scoring keys did not increase, to any practical extent, the accuracy of college attendance predictions made from the general key formed on the total group. In fact, the general key performed about as well as the special ability keys at each of the six ability levels. Taken together, the results from these two studies appear to indicate that the general nature of effective biodata items does not substantially change from one ability level to another.

The college attendance predictions based on the SIB responses of students in Subsample 4 were in the form of probabilities. Any student who had an estimated probability of attendance greater than .50 was classified as an attender. The remaining students were classified as nonattenders. When the predictions were compared with the actual status of the students, the hit rate was found to be 76%. The corresponding point biserial correlation coefficient was .58. The chance rate for correct predictions is the same as the college attendance rate for students in this study. Since this rate was about 51%, the obtained hit rate represents a 50% improvement over the chance rate. It would appear that the power of biodata as a predictor of college attendance is substantial

even when simple keying techniques and a grossly defined criterion are used.

Discussion

An optimally weighted combination of biodata and academic aptitude did not substantially improve upon the accuracy of predictions obtained from biodata alone. The hit rate for the combination was 77% with the corresponding point biserial coefficient being .60. On the other hand, the hit rate using only academic aptitude as the predictor was 71%, and the equivalent point biserial coefficient was .48. Thus, predictions based on biodata in combination with academic aptitude substantially improved upon the accuracy of predictions based on academic aptitude alone.

Some of the effectiveness of biodata as a predictor may be due to the inclusion of items in the academic background category. Eight of these items deal with student estimates of their high school grades and, hence, might be considered to be redundant if data on high school GPA were available. In order to determine the relationship between biodata and college attendance independent of GPA and other information on a student's academic background, a separate scoring key omitting items from the academic background category was developed. Since no effort was made to redetermine optimum key length, this scoring key was identical to the general key in every other respect. When the abbreviated scoring key was applied to the SIB responses of students in Subsample 4, the point biserial correlation between biodata scores and college attendance was found to be .52. The decrease from the value of .58 reported for the general key is substantial. However, the performance of the

abbreviated biodata key still compares favorably with the coefficient of .48 obtained by use of academic aptitude scores. When both variables were used in weighted combination, the multiple point biserial correlation coefficient was .57, almost as high as when the full set of biodata was used. Regardless of whether items in the academic background category are considered to be biographical in nature, it can be seen that biodata make an important contribution to the prediction of college attendance.

The percentages shown in the lefthand data column in Table 3 represent the base rates per biodata category to be expected if items meeting the cutoff criterion are no more likely to fall in one category than in any other. For example, one would expect, on the basis of chance alone, that 8% of the biodata items meeting the cutoff criterion would fall in the extracurricular category. Good sources of predictors should be found in biodata categories for which the percentage of items differentiating attenders and nonattenders across the various ability levels consistently exceeds the base rate. Perhaps the best examples would be the academic background and extracurricular categories. Most notable for lack of effectiveness would be the student health category. One should keep in mind, however, that the effectiveness of specific items is not indicated in Table 3. Detailed information on the performance of specific SIB items has been presented by Flanagan et al., (1964).

This study has demonstrated that, for students in general and for students falling in each of six ability level groups, many different types of biodata are predictive of college attendance. As a predictor of this specific criterion, biodata can best be viewed as reflecting psychological factors which influence college attendance. Biodata do

not, after all, make the choice of attending college for the student. Rather, they represent covariates or symptoms of intervening variables which are the real motivators and facilitators. It was not the purpose of this study to speculate on the manner in which biodata are related to these motivators or to other socially significant behavior. What has been shown is that biodata are potent and pervasive predictors of college attendance, predictors that are not merely concomitants of academic aptitude.

References

Beezer, R. H., & Hjelm, H. F. Factors related to college attendance. Cooperative Research Monograph, 1961, No. 8. (US HDEW Publ. No. OE-54023)

Berdie, R. F., & Hood, A. B. Decisions for tomorrow. Minneapolis: University of Minnesota Press, 1965.

Cooley, W. W., & Lohnes, P. R. Multivariate procedures for behavioral sciences. New York: Wiley, 1962.

Flanagan, J. C., & Cooley, W. W. Project TALENT one-year follow-up studies. Pittsburgh, Pa.: School of Education, University of Pittsburgh, 1966.

Flanagan, J. C., Davis, F. B., Dailey, J. T., Shaycoft, M. F., Orr, D. B., Goldberg, I., & Neyman, C. A., Jr. The American high-school student. Pittsburgh, Pa.: Project TALENT Office, University of Pittsburgh, 1964.

Heist, P. The entering college student--background and characteristics. Review of Educational Research, 1960, 30, 285-297.

Medsker, L. L., & Trent, J. W. The influence of different types of public higher institutions on college attendance from varying socioeconomic and ability levels. Berkeley, California: Center for the Study of Higher Education, 1965.

Prediger, D. J. Development of moderated scoring keys for psychological inventories. Educational & Psychological Measurement, 1969, 29(No. 4), in press.

Project TALENT. The Project TALENT Data Bank. Pittsburgh, Pa.: Project TALENT Office, University of Pittsburgh, 1965.

Footnotes

[1]The author is grateful for the use of data from the Project TALENT Data Bank, a cooperative effort of the United States Office of Education, the University of Pittsburgh, and the American Institutes for Research. The design and interpretation of the research reported herein, however, are solely the responsibility of the author.

The research reported herein was partially supported by a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

Table 1

Summary Description of Biographical Items

Item category and description	No. of items	SIB item No.[a]
Student activities		
Extracurricular: membership and offices held in organizations, dating experience, age when dating began, frequency of nights out	21	1-13, 51-55, 229-231
Hobbies and pastimes	25	14-33, 292-296
Work experiences: type, amount, source of funds	17	34-50
Reading: amount and type	9	56-64
Awards: subject matter areas, athletics, clubs and activities	8	284-291
Academic background		
Curricular tract, number of school transfers, attendance, hours spent studying, semesters spent in various curricular areas, grades in various curricular areas	23	91-113
Guidance		
Frequency, topic, source	16	114-129
Health of student		
Frequency and length of sickness, hours sleep, presence or absence of variety of physical defects or ailments	45	227, 228, 241-283

(Table continued on next page)

Table 1 (continued)

Item category and description	No. of items	SIB item No.[a]
Family and home		
Employment of parents: occupation of father and mother, main source of support, responsibilities amount of income	13	130-135, 173-175, 206, 208, 213, 214
Community activities: parents' membership and degree of involvement in organizations by type	20	138-157
Nature of dwelling: type, size, cost, number of appliances, sports and entertainment devices, number and age of car(s), etc.	16	169-172, 190-199, 225-226
Books and magazines: number of books in home, number and type of magazines	14	176-189
Education of parents and siblings: level, type	7	201-205, 218, 219
Miscellaneous: Number and birth order of siblings, age of father and mother, parents natural born?, parents both in home?, location of previous residence, etc.	12	136, 137, 167, 168, 200, 215, 216, 220-224
Total number of items	246	

[a]Corresponds to item number on Project TALENT Student Information Blank (Flanagan et al., 1964)

Table 2

Frequencies and College Attendance Rates by Ability Level

Ability level group (Subsample 1)	N	No. of attenders	% attending
1 (low)	1693	268	16
2	1697	503	30
3	1694	748	44
4	1706	977	57
5	1702	1238	73
6 (high)	1691	1448	86
Total subsample	10183	5182	51

Table 3

Percentages of Biodata Items Meeting Cut-off Criterion at Different Ability Levels

Biodata category	% of available items falling into category	Separate ability level groups						Total group
		1	2	3	4	5	6	
Student activities								
Extracurricular	8	16	14	12	12	14	9	10
Hobbies & pastimes	10	9	9	9	7	6	7	9
Work experience	7	3	4	4	5	4	9	5
Reading	4	9	6	4	6	4	0	3
Awards	3	0	0	3	3	0	4	4
Academic background	9	15	14	14	13	18	24	14
Guidance	7	15	7	9	7	6	8	7
Health of student	18	0	3	6	3	6	0	12
Family and home								
Employment of parents	5	5	6	6	7	8	4	7
Community activities	8	10	13	11	10	7	12	5
Nature of dwelling	7	6	8	7	11	10	4	9
Books & magazines	6	3	8	6	6	8	8	6
Education of parents	3	6	6	4	4	5	5	4
Miscellaneous	5	2	2	4	6	4	7	4

Note.--Columns should add to 100% except for the effects of rounding errors. The number of items meeting the cutoff criterion at ability levels 1 through 6 and for the total group were as follows: 67, 110, 110, 112, 100, 76, and 136.