DOCUMENT RESUME

ED 221 150 HE 015 546

AUTHOR Eagle, Norman
TITLE An Attempt to Optimize the Early Identification of the Potential Dropout: Use of Dynamic Predictors, Differentiated Criterion Groups, and Alternative Discriminant Function Methodology.
PUB DATE 17 May 82
NOTE 40p.; Paper presented at the Annual Forum of the Association for Institutional Research (22nd, Denver, CO, May 16-19, 1982).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; *College Freshmen; College Students; Discriminant Analysis; *Dropout Characteristics; Dropout Research; Grade Point Average; Higher Education; High Risk Students; Majors (Students); *Predictive Measurement; Predictor Variables; *Student Attrition

ABSTRACT

A system to identify the potential dropout was studied at a two-year urban institution. All entering freshmen participating in a compulsory freshmen orientation course were administered a biographical inventory during the fifth week of enrollment. An academic or scholastic index was derived, based on high school averages and basic skills placement scores in reading, writing, and mathematics. The spring 1981 group included 328 students, while the fall 1981 group involved 704 students. At the beginning of the second semester, all students were classified into one of four groups: dropout, failing index; dropout, passing index; returnee, failing index; or returnee, passing index. In addition, the spring 1981 cohort was reclassified at the beginning of the third semester. Each group was divided into those who had earned satisfactory academic indices and those who had not. Discriminant analyses were conducted using two, three, and four group criteria. After analyses were carried out on the total cohorts, the data were divided into four major curriculum areas: liberal arts, science and technology, nursing and medical laboratory technology, and business. A separate group of students was evaluated to cross-validate the proportion of correct group placements of known placements against a group of test students for whom placements were unknown. It is concluded that the system should not be implemented as it currently exists, and problems that prevented greater predictive success are briefly considered. Tentative findings suggest that the most important factors involved in student attrition may be related to ethnicity and language primacy, student commitment to educational goals, college workload, scholastic readiness, and ethnic group and marital status interactions. (SW)

An attempt to optimize the early identification of the potential dropout: use of dynamic predictors, differentiated criterion groups, and alternative discriminant function methodology.

Presented at:
Annual Forum
Association for Institutional Research
Denver, Colorado

May 17, 1982

Dr. NormanEagle, Director
Office of Institutional Research
Bronx Community College
Bronx, N.Y. 10453

An attempt to optimize the early identification of the potential dropout: use of dynamic predictors, differentiated criterion groups, and alternative discriminant function methodology.

Purpose:

This study derived from an attempt by a college administration to identify by mid-first semester, entering freshmen showing high probabilities of not returning to the college for the second semester. Once identified by this early warning system, the college intended to examine the protocols of entering students assigned to the 'potential dropout' category for clues about the conditions which may have led to the identification of the student as a potential dropout. Following this, the college intended to bring all available supportive resources to bear in an effort to encourage the student's continued enrollment, if this appeared to be in the best interests of the student.

Background:

The study of student attrition at the college level has been of interest to college faculty and administrators for some time. In the early 1950s there was considerable interest in the subject as revealed by the rash of studies which appeared at that time, i.e., Iffert, R.E. (1955), Fullmer, D.W. (1956), Slocum (1956), and Grace, H.A. (1957). The problem was mainly of an 'academic' interest until the mid-seventies when inflationary pressures began to strain college budgets and financial resources. Student 'counts' and FTEs became crucially important, leading to greater emphasis on institutional studies which might reveal why students were dropping out and what colleges could do to improve retention. This is illustrated at the institution sponsoring the current study where from 25%-30% of all students qualified to return from Spring to Fall semesters in recent years, failed to do so. The loss of financial support (state aid) for these attriting students amounted to millions of dollars.[1]

[1]Within the group of attriting students is a subgroup of unknown size who transfer to other institutions. This is a central problem in the design of many attrition studies, and will be discussed later.

Two general approaches to the problem of attrition have been used in recent years. Tinto (1975), for example, has asserted that longitudinal, path analysis, types studies are essential if the complex inter-play and interactions of factors leading to attrition are to be understood. The prevailing approach, however, possibly in reaction to the pressure to produce immediate ameliorative applications, is shown in the numerous attempts over the past decade to select significant predictors of attrition and to combine these in prediction equations in order to identify 'high risk' students. This is the basic approach used in the current study (in response to the practical need at the College to stem the tide of departing students). However, several refinements to the usual procedure of developing regression equations were introduced into the current study. These will be reported in the discussion of methods and procedures.

One of the most difficult and crucial problems researchers have faced in the field of attrition research is the conceptualization and definition of attrition itself. When can a student be considered a dropout? How should the criterion, dependent variable groups, be defined? The lack of resolution of this issue has undoubtedly resulted in 'fuzzy', over-lapping, undifferentiated criterion groups with subsequent failures to find significant relationships with predictors. It has also led to the impossibility of gaining an over-view of the attrition problem, of conducting productive meta-analyses of results.

Several approaches have been advanced to deal with this problem. Vaughn (1968) and Tinto (1975), for example, have advocated the separation of voluntary fron non-voluntary withdrawals. Empirical support for this was found by Green and Morlock (1978) who found that these two groups differed on a number of characteristics. However, Pantages and Creedon (1978) in their comprehensive review of the attrition research literature recommend that this approach be abandoned claiming that "it is less confusing in the long run to regard academic dismissal as an intervening variable that leads to withdrawal rather than as part of the dependent variable."

Pantages and Creedon (1978) propose a three group solution to the criterion problem, as follows: 1)students who graduate in the 'normal time', 2)students who dropout but reenroll at a later date and graduate from any college 3)permanent dropouts. While this may be an improvement over the simple dropout vs. retaince dichotomy, a number of problems with the Pantages and Creedon approach remain, namely: the question of 'normal time' for graduation by no means has a clear and unambiguous answer,

the ability to track transferring students is limited, and the concept of a 'permanent dropout' assumes a finite duration over which students will be tracked. On what basis will that period be decided?

Another criticism of past attrition research has been the predominant use of static demographic and academic predictors, such as gender, socio-economic status, and high school averages. In response to this many researchers have begun to investigate the interactions or the congruence between the needs and characteristics of students and the corresponding 'needs' and characteristics of institutions. Researchers utilizing this approach have been Tinto (1975), Starr, Betz, and Menne (1972), Flannery, et.al (1973), Alfred (1974), Holland (1969), Festinger (1962), and Cope and Hannah (1975). This problem will also be addressed in our discussion of methods and procedures.

With regard to the current state of knowledge concerning factors or conditions associated with 'dropping out' of college, the following represents our summary of the results of the dozens of studies reported by Pantages and Creedon in their comprehensive review:

A. Conditions which appear <u>important</u>:

1) high school academic achievement, class rank, S.A.T. scores.

2) scholarships and grants (but not loans)

3) the relationship between student needs and psychological characteristics, and the college environment.

B. Conditions which appear <u>marginally important</u>:

1) gender

2) parental education

3) study habits

4) motivational and educational goals

5) peer influence

6) individual personality characteristics

C. Conditions which appear unimportant:

1) age

2) socio-economic status, when controlled for G.P.A.

3) size of high school

4) distance of college from home

5) public vs. private high school

While Pantages and Creedon analysed the varying importance of the many variables which have been used to identify the potential dropout, they paid little attention to the overall effectiveness of the multiple regression equations derived to make those identifications. In point of fact the record has not been very impressive. For example, Brown (1975) found that the mean R^2 in 78 studies of attrition from 1970 to 1974 was only .259, though Capoor and Eagle (1977) showed that these R^2 could be considerably enhanced by using subgroups and interactions.

The current study while making use of subgroup anlyses and interactions, also makes an attempt at cross-validation which most previous studies have failed to attempt.

Method:

The general design of this study called for:

a) the collection of 'predictor' data for matriculated freshmen entering the college, during their fifth week of enrollment (timing will be explained later).

b) identifying the students returning for the second semester (and later for the third), and those who did not; then dividing each group into those who had earned satisfactory academic indices and those who did not.

c) conducting discriminant analyses using two, three, and four group criteria, and utilzing several methods for the selection of significant classification (predictor) variables, including various interaction terms.

d) replicating the procedure in c) for subgroups of several curriculums and for sex groups within some of these curriculum groups.

e) cross-validating the proportion of 'correct group placements' of known placements against a group of test students for whom placements were unknown, and who were also not included in the development of the discriminant functions themselves.

A. Data Collection

All matriculated freshmen enteing the College (a two year, urban institution) who were enrolled in a compulsory freshman orientation course, were administered a biographical inventory which provided data bout the following: educational aspiration, importance attributed to a college education, socio-economic status, study habits, financial pressure, smoking habits, recency of graduation, foreign language background, grade expectations, academic work load, hours of employment, college choice, curriculum choice, confidence in ability and goals, locus of control, need for achievement, need for dominance, need for autonomy, perceived interactions with faculty, perceived faculty concern for students, and institutional goal commitment. The latter three predictors were adapted from Terenzini, Lorang and Pascarella (1980), while the 'need' predictors were adapted from the Jackson Personality Research Form.

From other records data were collected on gender, ehtnic group and marital status. An academic or scholastic index was derived through z score combinations of high school averages and basic skills placement scores in reading, writing and mathematics.

The instrument was administered not upon admission but during the fifth week of enrollment in order to allow students an opportunity to gain an impression of the college enviornment, their counselors, administrators, and instructors. This was a necessary condition for minimally reliable estimates of the Terenzini-Pascarella Scale scores which are related to student perceptions of the college environment. Ideally a longer period of time should have been allowed before these scores were collected, but since the intent of the project was to develop a system for the early identification of the potential dropout, more time for the development of perceptions of the college environment could not be allowed.

For the group entering Spring 1981, 328 students were included in the study while for the Fall 1981 entering class 704 students were included. In each case, a 50% sample was obtained.

B. Classification Groups

At the beginning of the second semester for these cohorts, all students were classified into one of four groups: 1)drop-out, failing index; drop-out, passing index; returnee, failing index; returnee, passing index. In addition the Spring 1981 cohort was re-classified at the beginning of its 3rd semester (Spring, 1982) and appropriate criterion group adjustments made.

In the Spring 1981 cohort 19% of the students were identified as drop-outs, a proportion which was considered marginally sufficient for the statistical analysis. However, only 11% of the Fall 1981 cohort proved to be 'dropouts'. The combined proportion (14%) may, therefore, have been inadequate for a reliable estimate of the characteristics of the 'drop-out' group, especially after this group required further division by academic success for some of the analyses.

C. Analyses

The technique used for examining the relationships between the classification variables (predictors) and classification groups (criteria) was multiple discriminant function anlaysis.

Separate analyses were performed by varying the number of the criterion groups and combining these with varying methods of selecting the classification variables (introducing them into the discriminant function equations).

With regard to the criterion or classification groups, two, three, and four group criteria were used. In the two group case, students were classified as either dropouts or returnees. In the three group case they were classified into dropout-failing, dropout-passing, and returning. In the four group case they were classified as dropout-failing, dropout-passing, returnee-failing, returnee-passing.

With regard to the selection of classification variables, the following procedures were used:

1) Wilks's Lambda - takes into consideration both the differences between groups and the cohesiveness or homogeneity within groups. Produces largest overall multivariate F.

2) Mahalanobis D. - selects the variable which generates the greatest separation for the pair of groups which are closest at that step.

3) Minimizing residual variance MINIRESID)-selects the variable which produces the lowest residual variance between groups.

4) Rao's V - is a generalized measure of group separation and does not concern itself with cohesiveness within groups.

5) Maximum-Minumum F - produces the largest pairwise multivariate F for the two groups with the smallest F on that variable.

All of the above are step-wise procedures which used an F=1.5 as the value for inclusion in, or exclusion from, the analyses as the process of variable selection proceeded. In addition, the Direct method was used, which simply allowed all of the classification variables to enter into the analysis (thus rendering the order of inclusion irrelevant).

Dummy variables were used as classification variables in the cases of sex, ethnic group, and marital status, and interaction dummy groups were set up for the interaction between ethnic group and marital status. This was done because the size of the population was considered inadequate to support fragmentation beyond subdivision by curriculum (although a further subdivision by sex was attempted 'experimentally' in one or two cases).

The data from the Spring 1981 and Fall 1981 cohorts were merged to produce a single group of 1,032 students whose status at the beginning of their respective second semesters was used for group classification. Then the Spring 1981 cohort alone was examined for student dropout-retention status at the beginning of their third semester (end of first year), and a similar series of analyses was performed on these data.

D. Replication

After analyses were carried out on the total cohorts, the data were divided into four major curriculum areas: liberal arts (non science), science and technology, nursing and medical laboratory technology, and business. In some cases the liberal arts and business curriculums were further subdivided by sex, even though the small number in these groups precluded any attempt at cross-validation.

E. Cross-validation

While the college was, of course, interested in some of the general conditions and student characteristics relating to it own body of dropouts it was mostly interested in developing a system which could efficiently identify a potential drop-out early in the first semester. It was, therefore, mainly interested in the ability of the procedure to place students into their 'correct' criterion groups, or more specifically, into either a high dropout probability or high returnee probability group, rather than in the relative or unique importance of any of the classification variables or conditions themselves.

An important part of the S.P.S.S. discriminant function analysis is the assignment of each subject (student) to a criterion group to which the student has the greatest probability of belonging. In this study a small representative sample was withheld from the generation of the discriminant functions. This sample was used as the cross validation (test) sample. When this was done (for large subgroups), the predicted group placement was compared with the actual group placement of students in the test samples, yielding a 'cross-validation' of the procedure's effectiveness.

Results

Results of the various analyses, utilizing different criterion groupings and different classification variable selection procedures are shown in Tables 1-9.

Tables 1-3 deal with the 328 students in the Spring 1981 cohort after one semester only. Since these data were merged with the Fall 1981 cohort data for subsequent analyses, only one reference to these tables will be made. For example, Table 1 shows clearly that the Wilks's Lambda, Mahalanobis D, Rao V, and Minimum residual selection procedures produced identical results, both in terms of measures of relationships-separation, and in terms of the variables selected and standardized weights assigned these variables. Therefore, in subsequent analyses any one of these procedures is used to represent this entire block of selection criteria. Also, it is clear that the remaining two methods do not produce results very different from the four mentioned above.

Tables 4-7 deal with the data merged from both the Spring 1981 and Fall 1981 entering classes. The criterion here is the drop-out/return status at the beginning of the student's respective second semester.

Table 4 compares the use of two, three, and four group criteria for a single selection method, Maximum F (as mentioned before there is actually no clear superiority of any selection procedure method over the others). Table 4 suggests that the four group criterion may produce a slightly more 'efficient' analysis than the two or three group criteria. The first function canonical R is slightly larger (.34 vs. .24 and .25) and the amount of improvement over chance in the percent of correct placements of test subjects is greater than in the other two cases (55%-25% vs. 43%-33% and 63%-50%). On the other hand since the major interest lies in the total proportion of cases correctely identified, the two category case appears 'superior', even though its degree of gain over chance is the smallest.

Table 5 indicates clearly that there are no differences in predictive efficiency among the three selection procedures used. Any procedure used with the entire merged group of students produces between 65% and 67% correct placements of test subjects, against a chance correct placement of 33%. However, when business, nursing, and science curriculums are considered separately, an improvement in predictive efficiency is observed, from generally 65% to 75% correct group placements, against chance correct placement of 33%. For the two group criteria the table shows a 75% correct placement against a chance correct placement of 50%. It is important to notice the shrinkage in the proportion of correct placements due to cross-validation, i.e., from 86% to 65%.

When the four group criterion is used, and all variables are entered into the discriminant function, the maximum proportion of correct test placements is 63%, in the science and technology curriculum compared to chance correct placements of 25% (Table 6).

Finally, for the merged one semester data, Table 7 shows that no improvement is realized by using all the variables in the function where three criterion groups are used in the place of four. The highest proportion of correct placements is still 75% for the science and technology curriculum.

The table also suggests that the further division of curriculums into males and females has only a slight effect in the improvement of the percent of known cases classified correctly, and presumably even this slight improvement would experience some shrinkage in cross-validation. Group sizes were not sufficiently large to illustrate this through analysis.

With regard to the objective of secondary interest, the identification of some general characteristics and conditions relating to attrition, examination of the classification variables appearing most frequently for all curriculums combined in tables 4-7, and their standardized weights, suggests that the various discriminant function procedures tend to give emphasis to a group of perhaps seven or eight of the twenty-three original predictor variables.

What appears to be most important in separating the criterion groups is the combination of ethnicity and the language spoken in the home as a child (Figure 1). Commitment to an educational goal is also a predictor of modest significance and, among blacks, marital status also appears to contribute to the separation of criterion groups.

Finally, the number of credits taken by students (workload) and their scholastic readiness as measured by combining high school average with basic skills placement scores also appear to be contributing to the separation of the criterion groups.

When contributing predictors within specific curriculum groups are examined, some variations in the importance of variables are seen and some appear which did not appear under the combined curriculum condition. However, the size of these curriculum groups militates against a serious interpretation of the significance of these predictors. Indications are that different sets of predictors may be found effective within different curriculum groups.

The results of the corresponding analyses (over-all curriculums), but using beginning third semester dropout or returnee status to define the criterion groups are shown in Tables 8 and 9. These tables are of interest only as they may suggest a possible improvement in predictive power of the discriminant functions as the time period allowed for defining dropout status is extended. The comparisons between Tables 1 and 2, on the one hand, and Table 8 on the other, must be made tenuously, bearing in mind that the one year criterion cohort consists of only the 328 students entering as freshmen in Spring, 1981.

Examination of the tables reveals that no clear tendency for an improvement in predictability can be seen. For example, using the Direct method with the four group criterion, 30% of the one semester test cases are identified correctly compared to 33% in the two semester situation. However, using the

Wilks or the Maxim F selection procedures, the one semester condition appears slightly superior to the two semester condition (37% correct test identification vs. 33%).

Discussion and Conclusions

At best the results of this attempt to develop a system for the early identification of potential drop-outs have been inconclusive. The improvement achieved in prediction, while in some cases considerably greater than chance (75% vs. 33%) does not appear great enough, nor consistent enough, to warrant the system's operationalization at its current state of development.

We consider that three major problems have contributed to this. First, little progress was made in the solution of the 'criterion problem' by differentiating four criterion groups from two. Apparently this was insufficient in creating clearly differentiated criterion groups. Particularly at the kind of urban community college at which this project was carried out, where student movement in and out of enrollment is pervasive, this simple differentation appears inadequate. At different points in time a given student may be considered a dropout at one point and a returnee at another. In this type of setting looking for the 'dropout' may be illusory. Since high proportions of students graduating from this institution 'stop-out' many times before graduating, it may be better to reformulate the research problem as one of identifying the potential graduate or non graduate over a long term, i.e., eight or ten semester period, (at this 'two-year' college) rather than the identification of the short term dropout. In this connection it is also imperative, for the formation of maximally differentiated criterion groups, to account for transfers. Clearly a significant proportion of the students considered as dropouts or non returnees in this project had enrolled at other institutions and were misclassified in the analysis.

The second problem militating against greater predictive success was the lack of balance representated in the sample betwen drop-outs and returnees. Before the project began approximately 25% of the students in any entering class did not return for the second semester. Whether reacting to a strong administrative program aimed at improving retention, or because of bias in the sample, only 19% of the Spring 1981 cohort and only 11% of the Fall 1981 cohort could be included in the drop-out grour Even when both cohorts were combined for analysis this decline in the proportion of dropouts produced a relatively small drop-out group

(less than 140 after test cases were removed), probably inadequate for acceptable reliability. This weakness could be mitigated by the collection of more data, at least twice as much as was available for the project and, if separate curriculum group analyses are also to be conducted (the most promising approach), at least six or seven times as much data will be required (the rule applied is that in any classification group there should be ten times as many subjects as the total number of predictor variables).

Finally, there is the choice of the classification variable battery, and the reliability of the measures used. Although variables were selected to represent those reported in the literature as being effective in identifying potential dropouts, it is very possible that the best mix of predictors was not assembled, and that the most effective inter-action terms were not utilized. Certainly only a very small fraction of the total number of possible interactions was explored in this project. Also, several of the predictor measures were derived from only two or three items (i.e., S.E.S., educational importance) so that the reliability of such measures is in question. Even measures drawn from empirically supported scales (i.e., Jackson PRF, Rotter's Locus of Control, Terenzini's integration scales), the adaptations used for this study, and their applications for the specific student population used, raise further questions of reliability.

The strongest conclusions which can be reached from the results obtained must be very tentative at best. Bearing this in mind, it might be said that the following tentative trends have been suggested by the findings, for the unique student body at the institution in question: the most important factors involved in dropping out may be related to ethnicity and language primacy, ethnic group-marital status interactions, student commitment to educational goals, student college work load, and scholastic readiness.

Only with the adoption of the refinements in design mentioned above would it be possible to test the actual importance of these suggested trends and to establish the feasibility of developing a workable system for the identification of the potential non-returnee or non-graduate at the college at which this study was carried out.

Table 1. Comparison among six classification variable selection procedures.
One semester analysis - Spring 1981 cohort - Four group case.
Total data, N = 328

Selection Procedure	eigenvalue % 1st function	1st func. R	% known cases classified correctly	% test data (n=30) classified correctly vs. expected		classification variables selected, and standardized weights on 1st disscriminant f.	
DIRECT	69%	.53	63%	30%	(25%)	BS .68 HISPANIC -.65 BLACK -.59 SINDEX .57	HS .53 SINGLE .53 LANG .37 GRD EXP .37
WILKS LAMBDA	74%	.47	62%	37%	(25%)	LANG .55 SINDEX .54 GRD EXP -.42 WKHRS -.36	NAUT -.34 FCNCRN .31 HISPANIC -.21 NDOM .26
MAHALANOBIS D	74%	.47	62%	37%	(25%)	LANG .55 SINDEX .55 GRD EXP -.42 WKHRS -.36	NAUT -.34 FCNCRN .31 HISPANIC -.27 NDOM .26
RAO V	74%	.47	62%	37%	(25%)	same	
MINIRESID	74%	.47	62%	37%	(25%)	same	
MAXMIN F	73%	.48	64%	37%	(25%)	SINDEX .56 LANG .48 GRD EXP -.39 WKHRS -.36	NAUT -.36 HISPANIC -.33 NDOM .29 FCNCRN .26

Table 2. Comparison among three dependent variable groupings.*
One semester analysis - Spring 1981 cohort - Wilks selection criterion. Total data, N = 328

Criterion	eigenvalue % 1st function	1st func R	% known cases classified correctly	% test data (n=30) classified correctly vs. expected		classification variables selected, and standardized weights on discriminant f...	
Two Group	100%	.30	81%	67%	(50%)	INTGLS .45 SINDEX .39 BCCLOD .39	WKHRS -.38 HM -.35 ACHMOT -.30
Three Group	88%	.34	81%	67%	(33%)	LANG .66 HISPANIC -.61 INTGLS .47	SINDEX .45 HM -.38 WKHRS -.37
Four Group	74%	.47	62%	37%	(25%)	LANG .55 SINDEX .54 GRD EXP -.42	WKHRS -.36 NAUT -.34 FCNCRN .31

* Two group criterion = returnees, non returnees

Three group criterion = returnees, non returnees (passing), non returnees (failing)

Four group criterion = returnees (failing), returnees (passing), non returnees (failing), non returnees (passing).

Table 3. Comparisons by curriculum groups.
One semester analysis - Spring 1981 cohort,
Wilks criterion.

Curriculum Group	N	Criterion Group	1st func R	eigenvalue % 1st function	% known cases classified correctly	classification variables selected, and standardized weights on 1st function.	
Science & Tech	70	2	.63	100%	83%	BLACK .75	INTGLS .55
						EDIMP -.70	LANG .47
						CONFID -.63	SINDEX .47
"	70	3	.56	60%	79%	BLACK 1.20	LANG .42
						HISPANIC .74	EDIMP -.40
						FACINT .48	NACH .35
"	70	4	.62	44%	67%	HISPANIC 1.07	CURCHS .69
						BLACK .89	FCNRN .88
						SMOK .70	SEX .55
Data Processing	46	2	.49	100%	85%	YRGRAD .65	LOCUS .47
						SINGLE .59	STHAB -.42
						INTGLS .50	
"	46	3	.36	64%	81%	YRGRAD .96	LANG .67
"	46	4	.80	81%	79%	SINDEX 1.03	FACINT .73
						YRGRAD .76	NACH -.60
						LANG .74	INTGLS .39

Table 3 (Continued) Comparisons by curriculum groups. One semester analysis - Spring 1981 cohort, Wilks criterion.

Curriculum Group	N	Criterion Group	1st func R	eigenvalue % 1st function	% known cases classified correctly	classification variables selected, and standardized weights on 1st function.
Nursing & Med Lab	86	2	.57	100%	85%	INTGLS .70 SEX .53 YRGRAD -.58 LANG .45 NAUT -.55 SINGLE .41
"	86	3	.53	57%	89%	YRGRAD -.77 BM .47 LANG .66 NAUT -.39 INTGLS .65 MARRIED -.33
"	86	4	.48	52%	64%	MARRIED 1.03 EDIMP .57 SINGLE .81 BLACK .12 HM -.75 NAUT -.07

Table 3. (Continued) Comparisons by curriculum groups. One semester analysis - Spring 1981 cohort, Wilks criterion.

Curriculum Group	N	Criterion Group	1st func R	eigenvalue % 1st function	% known cases classified correctly	classification variables selected, and standardized weights on 1st function.
Business	73	2	.52	100%	84%	BCCLOD .59 WKHRS -.38 NAUT .54 BS -.33 SINDEX .52
"	74	3	.50	54%	88%	SMOK -.83 NAUT .41 N DOM .51 BCCLOD .29 EDPAR .50 SINDEX -.16
"	74	4	.69	66%	74%	SINDEX .74 FCNCRN .38 GRDEXP -.66 BCCLOD .32 WKHRS -.48 NAUT -.25
Business (Males)	36	4	.92	66%	92%	YRGRAD -1.30 CURCHS .93 WKHRS 1.08 SINGLE -.71 BCCCHS .95 SINDEX -.64
Business (Females)	39	4	.69	61%	82%	GRDEXP-.81 SINDEX .27 NDOM .49 NAUT .09 BCCLOD .47

Table 4. Comparisons among three dependent variable groups. One semester analysis combining Spring 1981 and Fall 1981 cohorts, using Max F for selection of variables.

N=1,032 Test cases = 40

Criterion n of groups	eigenvalue % 1st func	1st func R	% known cases classified correctly	% test data correct vs. expected	classification variables selected, and standardized weights on 1st function.
2	100%	.24	86%	63% (50%)	LANG .56 BCGLOD .39 INTGLS .49 BM .31 SMOK −.47 BLACK .28
3	72%	.25	86%	65% (33%)	LANG .53 BM .38 HISPANIC −.49 BCGLOD .35 INTGLS .43 SMOK −.30
4	61%	.34	64%	55% (25%)	SINDEX .60 GRDEXP −.46 LANG .47 WKHRS −.31 HISPANIC −.46 STHAB .26

Table 5. Comparisons among selection methods, one semester analysis, for three criterion groups, combining Spring 1981 and Fall 1981 cohorts.

Method	N	eigenvalue % 1st function	1st func R	% known cases classified correctly	% test cases correct (40)-(16)	classification variables selected, and standardized weights on 1st function.
DIRECT	1,032	69%	.27	86%	67% (33%)	BS .66 BM .44 HISPANIC -.51 INTGLS .44 LANG .48 BLACK -.36
MAX F	1,032	72%	.25	86%	65% (33%)	LANG .53 BM .38 HISPANIC -.49 BCCLOD .35 INTGLS .43 SINGLE .29
WILKS	1,032	72%	.25	86%	65% (33%)	same
Business (2 groups)	187	61% -	.42 (.32)	87% (85%)	75% (33%) (75%)(50%)	BS .93 BLACK -.54 BM .72 CONFID .39 LANG .71 SEX .30
Nursing (2 groups)	266	76% -	.38 (.39)	84% (86%)	75% (33%) (75%)(50%)	SMOK -.64 LOCUS .41 SINDEX -.62 GRDEXP .33 MARRIED .49 EDIMP .28
Science (2 groups)	211	73% -	.45 (.41)	85% (86%)	75% (33%) (75%)(50%)	BLACK .80 EDIMP -.51 EDPAR -.75 STHAB .46 HS .60 BCCLOD .38

Table 6. Comparisons among curriculum subgroups for method Direct, 4 group criterion, one semester analysis, combining Spring 1981 and Fall 1981 cohorts, Test cases=40.

Curriculum	N	1st func R	eigenvalue % 1st function	% known cases classified correctly	% test cases correct	classification variables selected, and standardized weights, 1st function	
All Curriculums	1063	.34	60%	63%	60% (25%)	SINDEX .63 BS .43 GRDEXP -.41	LANG .37 HISPANIC -.34 WKHRS -.31
Business	187	.57	62%	72%	50% (25%)	HM 1.03 SINDEX -.77 BLACK -.66	MALE .59 BM .44 LANG -.42
Nursing	266	.46	39%	72%	44% (25%)	BS .97 LANG .64 BLACK -.51	BM .50 HM -.40 SINDEX .33
Science	211	.56	52%	71%	63% (25%)	HS -1.9 HISPANIC 1.67 SINGLE 1.48	BS -1.22 MALE .97 HM -.93

Table 7. Analysis by curriculum groups and sex after one semester, using combined Spring 1981 and Fall 1981 cohorts, three classification groups, and Direct variable selection method.

Curriculum	N	1st func R	eigenvalue % 1st function	% known cases classified correctly	% (16) test cases correct	classification variables selected, and standardized weights, 1st function.
All Combined	1032	.27	69%	86%	67% (33%)	BS .66 BM .44 HISPANIC -.50 INTGLS .44 LANG .48 BLACK -.36
Nursing	282	.43	53%	88%	67% (33%)	BS .83 BLACK -.51 BM .64 HM -.35 LANG .53 CONFID .29
Science and Technology	231	.50	75%	85%	75% (33%)	HISPANIC 1.75 BS -1.18 HS -1.72 MARRIED 1.14 SINGLE 1.39 HM -.86

Table 7 (continued) Analysis by curriculum groups and sex after one semester using combined Spring 1981 and Fall 1981 cohorts, three classification groups, and Direct variable selection method.

Curriculum	N	1st func R	eigenvalue % 1st function	% known cases classified correctly	% (16) test cases correct	classification variables selected, and standardized weights, 1st function
Business (M & F)	202	.42	62%	85%	69% (33%)	HISPANIC -1.22 HS 1.04 BLACK -1.13 SINDEX -.75 BS 1.12 SINGLE -.56
Males	94	.58	60%	91%		
Females	108	.60	74%	89%		
Data Processing (M & F)	140	.47	60%	92%		LANG 1.09 EDPAR .39 BLACK .73 STHAB -.36 YRGRAD .48 FACINT -.35
Males	50	.73	62%	94%		
Females	90	.58	65%	93%		

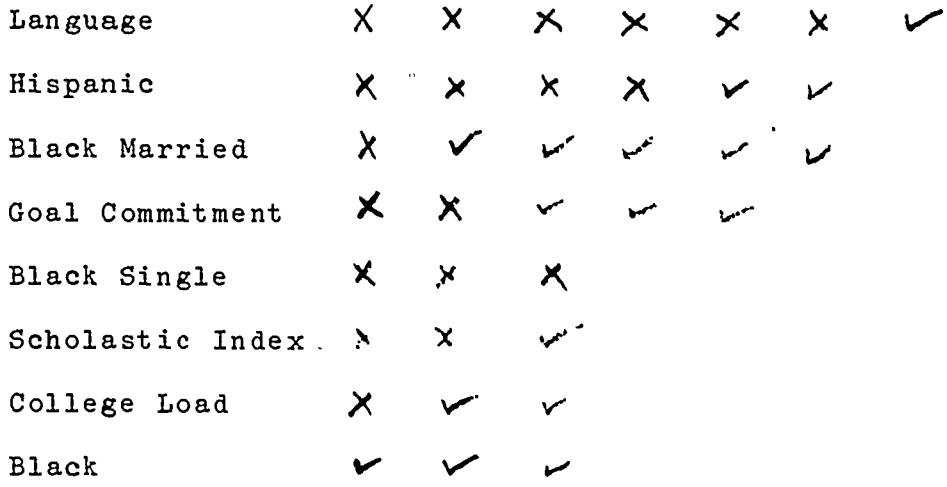
Table 8. One year criterion, students entering Spring 1981 only, over all curriculums. (N=328)

Groups	Method	eigenvalue % 1st function	1st func R	% known cases classified correctly	% test cases classified correctly	classification variable selected and standardized weights on 1st discriminant f.					
2	DIRECT	100%	.33	71%	63% (50%)	LANG	.57	YRGRAD	-.49	FACINT	.41
(N=329)						MALE	-.53	BLACK	.49	BCCLOD	-.34 N)
	WILKS	100%	.28	73%	57% (50%)	BLACK	.69	LANG	.52	WKHRS	-.44
						FACINT	.56	YRGRAD	-.51	SINDEX	.42
	MAXMIN F	100%	.28	73%	60% (50%)	(Same)					
3											
(N=329)	DIRECT	67%	.42	70%	53% (33%)	HISP	-1.02	SINDEX	.61	HM	-.50
						HS	.66	MALE	.61	STHAB	.32)
	WILKS	63%	.34	69%	63% (33%)	MALE	.93	SINDEX	.70	NAUT	-.32
						HM	-.80	HM	-.48	NACH	.31
	MAXMIN F	66%	.32	69%	63% (33%)	MALE	1.13	SINDEX	.63	SINGLE	.45
						HM	-.78	HM	-.46	NACH	.31
4											
(N=329)	DIRECT	61%	.54	56%	33% (25%)	SINDEX	.61	HM	-.51	HS	.44
						HISPANIC	-.59	LANG	.45	NAUT	-.29
	WILKS	69%	.51	56%	33% (25%)	SINDEX	.62	GRDEXP	-.34	HISPANIC	-.29
						LANG	.49	NACH	.32	NAUT	-.28
	MAXMIN F	68%	.51	57%	33% (25%)	(Same)					

Table 9. One year criterion for three curriculum groups, Spring 1981 students only. (N=328).

Curriculum	Method	N	1st func R	eigenvaluc % 1st function	% known cases classified correctly
Business	WILKS	71	.47	72%	74%
	DIRECT		.75	74%	90%
Nursing	WILKS	88	.61	68%	84%
	DIRECT		.57	57%	78%
Scienc•	WILKS	76	.61	61%	76%
	DIRECT		.73	63%	88%

Figure 1. Number of times predictors appeared among highest three or highest six in eight analyses involving combined curriculums, one semester criterion. (✗ among highest 3, ✓among highest 6).



Appendix A

Classification Variable Labels

Item #	Label	
4-5	ACHMOT	Educational aspiration or motivation
6	BCCCHS	Was BCC first choice of college within CUNY?
14	BCCLOD	Credit load taken by student
25-27	CONFID	Confidence in ability and goals
2	CURCHS	Is student in curriculum of first choice?
10	EDIMP	Perceived importance of education
11-12	EDPAR	Parental education
19-20	FPRESS	Financial pressure
24	GRDEXP	Expected grades at BCC
23	LANG	Main language spoken in house as a child
28-32	LOCUS	Internal-external locus of control
3	SEX	Male, Female
21	SMOK	Amount of cigarettes smoked per week
16-18	STHAB	Study habits
7-8	WKHRS	Number of hours of empolyment per week
22	YRGRAD	Year of graduation from high school
p3-4	NACH	Need for achievement
	NDOM	Need for dominance
	NAUT	Need for autonomy
p4	FACINT	Perceived interactions with faculty
	FCNCRN	Perceived faculty concern for students
	INTGLS	Institutional and goal commitment
	SINDEX	Index of high school average and basic skill placement test scores
	HM	Hispanic Married
	BM	Black Married
	HS	Hispanic Single
	BS	Black Single
		Marital Status
		Ethnic Group (Black, Hispanic, Other)

References

Alfred, R.L. A conceptual investigation of student attrition in the comprehensive community college. Paper presented at annual meeting of the A.E.R.A., Chicago, April, 1974.

Brown, D.J. Down with the lineal model. American Ed. Research J., 1975, 72, 491-505.

Capoor, M. and Eagle, N. Dropouts in two year colleges: Better prediction with the use of moderator subgroups. Community/Junior College J., 1977, 2, 171-177.

Cope, R.G. and Hannah, W. Revolving College Doors: The Causes and Consequences of Dropping Out, Stopping Out, and Transferring. New York: John Wiley & Sons, 1975.

Festinger, L. A Theory of Cognitive Dissonance. Stanford: Stanford Univ. Press, 1962.

Flannery, J.; Asbury, C.; Clark C.; Eubanks, D.; Kerchval, B.; Lasak, J.; McWorther, J.; Skellings, L.; Smith, D.; and Sutton, C. Final report for the ad hoc committee to study attrition at Miami-Dade Community College, 1973. ERIC Document ED 085052.

Fullmer, D.W. Success and perseverance of university students. J. Higher Ed., 1956, 27, 445-447.

Grace, H. A. Personality factors and college attrition. Peabody J. of Higher Ed., 1957, 35, 36-40.

Green, D.M. and Morlock, J.V. Identifying the potential dropout. Paper presented at annual conference North East Association for Institutional Research, Oct., 1978, University Park, Pa.

Holland, J.L. The Psychology of Vocational Choice: A theory of Personality Types and Model Environments. Waltham, Mass: Blaisdell, 1966.

Iffert, R.E. The student retention and withdrawal study. College and University, 1955, 30, 406-411.

Pantages, T.J. and Creedon, C.F. Studies of college attrition: 1950-1975. Rev. of Ed. Research, 1978, 48, 49-101.

Slocum, W.L. Social factors involved in academic mortality. College and University, 1956, 32, 453-64.

References continued

Starr, A.; Betz, E.L. and Menne, J. Differences in college student satisfaction: Academic Dropouts, non-academic dropouts. J. of Counseling Psych. 1972, 19, 318-322.

Terenzini, P.T.; Lorang, W.G.; and Pascarella, E.T. Predicting freshman persistance and voluntary dropout decisions: a replication. Paper presented at the annual conference of the American Educational Research Association, Boston, April, 1980.

Tinto, V. Dropouts from higher education: A theoretical synthesis of recent research. Review of Ed. Research, 1975, 45, 89-125.

Vaughan, R. College dropouts: Dismissed vs. withdrew. Personnel and Guidance J. , 1968, 46, 685-689.