Session 2793

Identifying Factors Influencing Engineering Student Graduation and Retention: A Longitudinal and Cross-Institutional Study

Guili Zhang, Tim Anderson, Matthew Ohland, Rufus Carter and Brian Thorndyke
Educational Psychology Department, University of Florida / Department of Chemical Engineering, University of Florida / Department of General Engineering, Clemson University / Educational Psychology Department, University of Florida / Department of Physics, University of Florida

Abstract
In this study, pre-existing factors are quantitatively evaluated as to their influence on student success. This study uses a database of all engineering students in the time period 1987 through 2000 and considers two definitions of success. The first, graduation, is defined as graduation in an engineering degree program as of the latest year in the study. The second, retention, is defined as either graduation or current enrolment in an engineering degree program as of the latest year in the study. A multiple logistic regression model was formulated to test for and estimate the predictive relationships between these measures of success and a set of six background variables that represent student's pre-existing demographic and academic characteristics (gender, ethnicity, high school GPA, SAT math score, SAT verbal score, and citizenship status). It is found that both graduation and retention in engineering for students who enter in an engineering discipline depends significantly upon high school GPA and math SAT scores, while verbal SAT scores correlated negatively with odds of graduation for five out of eight universities. Gender, ethnicity and citizenship also showed significant effects for some Universities, but these were not consistently positive or negative predictors. We also find that gender, verbal SAT scores, ethnicity and citizenship frequently appear as predictors of retention, but not as predictors of graduation.

Introduction
Identifying those factors that influence retention should be useful in suggesting approaches to improving student success in engineering. The identification of these factors will assist in developing meaningful admission procedures as well as aid the counseling and advising of students seeking an engineering degree. Much research has focused on identifying predictors of success in college and in engineering. Astin's 1965 study of 36,581 students indicated that the student's academic record in high school was the best single indicator of how well they would do in college[1]. He also indicated that there was a clear positive relationship between students' performance on tests of academic ability (e.g. SAT) and performance in college. Astin also listed gender as useful in predicting college freshman GPA. In a more recent study, Seymour and Hewitt[2] reported that the students leaving engineering were academically no different than those that remained. They reported students left for reasons relating to perceptions of the institutional culture and career aspects.

Perceptions and attitudes of engineering students have been examined in the literature. Besterfield-Sacre, Moreno, Shuman and Atman developed the Pittsburgh Freshman Engineering

Attitude Survey (PFEAS)[3]. They administered the survey at the beginning of the students first semester and again at the end of the first semester or the end of the first academic year. They report gender differences for female engineering students on the pre-survey. Female engineering students began their engineering programs with lower confidence in background knowledge about engineering, their abilities to succeed in engineering, and their perceptions of how engineers contribute to society than their male counterparts[3]. Those same female students indicated they were more comfortable with their study habits than did the male students. Differences for minority students were reported for African American vs. majority students, Hispanic vs. majority and Asian Pacific vs. majority students.

Zhang and RiCharde examined 462 freshmen that matriculated in the fall of 1997. Roughly 32% of these students were engineering majors[4]. They tested several cognitive, affective, and psychomotor variables to see which were significant predictors of college persistence. Their logistic regression identified self-efficacy and physical fitness as positive predictors of freshman retention, while judgment and empathy were negatively associated with persistence. They reported three reasons for freshman attrition: inability to handle stress, mismatch between personal expectations and college reality, and lack of personal commitment to a college education.

Levin and Wyckoff gathered data on 1043 entering freshmen in the College of Engineering at Pennsylvania State university[5]. They developed 3 models to predict sophomore persistence and success at the pre-enrollment stage, freshman year, and sophomore year. Eleven intellective and 9 non-intellective variables were measured. For the pre-enrollment model, the variables best predicting success were high school GPA, Algebra score, gender, non-science points, chemistry score, and reason for choosing engineering. The freshman year model identified the best predictors of retention as grades in Physics I, Calculus I and Chemistry I. In the sophomore year model the best predictors of retention were grades in Calculus II, Physics II and Physics I. They noted that predictors of retention were dependent on the students' point of progress through the first 2 years of an engineering program.

Other studies indicate the freshman year is critical. Lebold and Ward indicated the best predictors of engineering persistence were the first and second semester college grades and cumulative GPA[6]. They also reported that students' self-perceptions of math, science and problem-solving abilities were strong predictors of engineering persistence.

In this study, over 10 years of data for 8 colleges of engineering in 9 universities were used to evaluate pre-existing factors' influence on retention. Many studies have examined retention of engineering students for only one or two years. This snapshot approach while immediately informative does not offer the power of examining predictors over time. The cross-institutional nature allows us to compare the results across the universities to find their generalizability. Specifically, the 9 Universities are each public, but exhibit a wide range in other characteristics such as mission, minority and total enrollment, research emphasis, on-campus enrollment, and number of in-state residents enrolled. The longitudinal nature of our data allows us to look at change across time. Multiple logistical regression techniques allow us to examine the effect of each predictor while controlling for the other variables. We measured retention as both graduation and persistence in the engineering program. Our study looks at differences in predictors in these two definitions of retention.

Data Collection

This study uses the Southeastern University and College Coalition for Engineering Education (SUCCEED) longitudinal database (LDB) to identify pre-college entrance demographic and academic factors that predict engineering students' graduation. The LDB contains data from eight colleges of engineering involving nine universities: Clemson University, Florida A&M University, Florida State University, Georgia Institute of Technology, North Carolina A&T State University, North Carolina State University, University of Florida, University of North Carolina at Charlotte and Virginia Polytechnic Institute and State University. To protect the rights of human subjects, each university is assigned a letter that is only known by the researchers involved in the study. Throughout the paper, we examine the effects of predictors on two definitions of retention. For both definitions, we refer to the period 1987 through 1998, 1999 or 2000, depending on the latest year available in the LDB for a given institution.

In one set of analyses, retention refers to graduation in an engineering program during that time period, which we label *graduation*. Because it typically takes a student a minimum of four years to graduate, students who have entered university after 1995 have not usually had enough time to graduate, and are excluded from these analyses. Therefore, for the graduation analyses, we only include students matriculated in an engineering field between 1987 and 1994. The number of students used in the retention analyses are listed in the header of Table 1.G.

University	Cohorts	Graduation Percentage	Graduation Date
A	1987-1994	30.49%	1987-1998
B	1987-1994	24.50%	1987-1998
C	1987-1994	28.20%	1987-1998
D	1987-1994	35.54%	1987-1999
E	1987-1994	50.97%	1987-1999
F	1987-1994	54.33%	1987-1998
G	1987-1994	42.83%	1987-2000
H	1987-1994	43.04%	1987-2000
I	1987-1994	32.71%	1987-1999

Table 1.G. *Graduation data by university. Number of engineering students included in the analysis in descending order and not correlated to the alphabetic University designation: 11,382, 8,418, 7,072, 5,815, 2,542, 1,737, 1,065, 705, and 541.*

The second set of analyses defines retention as either graduation within that time period or current enrolment in the last year of the LDB, which we simply label *retention*. Thus retention analyses include students who have matriculated in an engineering field during any year from

1987 through 1998, 1999 or 2000. The number of students used in the retention analyses are listed in Table 1.R.

University	Cohorts	Retention Percentage	Retention Period
A	1987-1998	52.49%	1987-1998
B	1987-1998	36.97%	1987-1998
C	1987-1998	43.89%	1987-1998
D	1987-1999	46.42%	1987-1999
E	1987-1999	57.30%	1987-1999
F	1987-1998	64.68%	1987-1998
G	1987-2000	48.75%	1987-2000
H	1987-2000	58.86%	1987-2000
I	1987-1999	46.14%	1987-1999

Table 1.R. *Retention data by university. Number of engineering students included in the analysis in descending order and not correlated to the alphabetic University designation: 15,079, 12,928, 11,842, 7,574, 4,146, 2,619, 1,501, 1,004, and 856.*

We study the dependence of graduation and retention on six independent variables (or predictors): ethnicity (ETHNIC), gender (GENDER), high school Grade Point Average (HSGPA), SAT math score (SATM), SAT verbal score (SATV), and citizenship status (CITIZEN). HSGPA, SATM, and SATV are continuous numerical variables, while ETHNIC, GENDER, and CITIZEN are categorical variables having several levels. Specifically, ETHNIC has six levels: African American (AfrAm), Asian (Asian), Hispanic (Hisp), Native American (NatAm), White (White) and other (Other). GENDER has two levels: male (Male) and female (Female). CITIZEN is divided among three levels: U.S. citizen (Citizen), U.S. resident but not citizen (ResAlien) and foreign (NRAlien).

Pair-wise deletion is used wherever there is missing data. In essence, any student who has a missing value on any of the predictors is excluded from the study. For most institutions, this exclusion has minimal impact on the analysis. However, a serious missing value issue involves three universities in particular. The LDB does not contain high school GPA information for two of the universities, and the analyses on these two universities are done without the high school GPA predictor. In addition, one of the universities does not have SAT math, SAT verbal and high school GPA, and the analyses on that university are done with only GENDER, ETHNIC and CITIZEN as predictors.

Statistical Methods

Table 1.G lists the number of engineering students matriculated in the time period 1987-1994 and the number of students graduated in an engineering field as of the latest record in the database. For such data, we want to investigate whether a student's graduation likelihood can be predicted by certain factors. Since graduation has two outcomes, graduated or not graduated, a

logistic regression model is appropriate. Furthermore, because we want to test the significance of more than one predictor, a multiple logistic regression model is in order. Such a model allows one to test for each predictor's significance while controlling other predictors. A similar analysis is done for retention.

The general multiple logistic regression model is,

$$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_i X_i + \varepsilon ,$$

where Z is a dichotomous variable (Z = 1 represents success, while Z = 0 represents failure). X_1 - X_i are the predictors of Z. Whether a specific predictor X_i significantly predicts Z with other predictors controlled can be determined by testing the parameter β_i.

In our study, the analysis is conducted for each individual university separately. Using SAS version 8.1, a multiple logistic regression model was formulated to test for and estimate the predictive relationships (the parameter/slopes) between graduation, and the predictors GENDER, ETHNIC, HSGPA, SATM, SATV and CITIZEN.

Type III analyses of effects provide the magnitude of each predictor's effect by controlling the other predictors. In other words, the Type III effect can "strip off" the effect of other predictors and focus on the predictor under investigation. The Wald Chi-Squared statistics on the predictors' effects are reported along with a p-value. The Chi-Squared test of independence, proposed by Karl Pearson in 1900, is one of the common approaches to investigating statistical dependence[7]. It tests the null hypothesis: graduation is independent of the predictor. A large Chi-Squared statistic (which corresponds to a smaller p-value) provides evidence that the null hypothesis is false. Generally a p-value smaller than .05 is required to reject the null hypothesis. The Wald Chi-Squared statistics and p-values for graduation are reported in Table 3.G, while results for retention are listed in Table 3.R.

The Stepwise Selection Procedure is used to select predictors that effectively predict graduation. At each step, the Stepwise Selection Procedure selects the variable that has the strongest effect among the variables that have not entered the model. This process is repeated until no additional effect satisfies the 0.05 significance level criterion for entry into the model. The Chi-Squared statistics of variables that are selected by the Stepwise Selection Procedure are boldfaced in Tables 3.G and 3.R. If a given variable is significant to either graduation or retention, but not both, its Chi-Squared statistics are indicated by a single asterisk ("*"). If the variable is significant to both graduation and retention, its Chi-Squared statistics are indicated by a double asterisk ("**").

The β parameters (slopes) are estimated using Maximum Likelihood Estimates. With these estimated β parameter values, we can obtain the estimates of the Odds Ratio[7]. The estimated Odds Ratios are reported in SAS output and are based on maximum likelihood estimates as well. To understand the meaning of the Odds Ratio, consider the graduation analyses. For a continuous variable, an Odds Ratio provides the relative probability of graduation with one unit increase in the predictor. For example, for university A, the odds ratio estimate of graduation due to HSGPA is 3.634. This says that a given student is 3.634 times as likely to graduate as another student whose high school GPA is 1 point lower (e.g., 3.50 vs. 2.50 high school GPA). When the predictor is a categorical variable, the Odds Ratio is the ratio of probability of graduation between two levels on the categorical variable. For example, for university A, the Odds Ratio estimate for GENDER (female vs. male) is 1.341. It tells us that a female is 1.341 times as likely to graduate as a male. A 95% Wald Confidence Interval (CI) is provided for every Odds Ratio estimate. If the Wald CI *does not* include 1.0, then the probability of graduation *is* significantly different for the levels compared. If the Wald CI *does* contain 1.0, then the probability of graduation *is not* significantly different. The Odds Ratio analyses are similarly done for retention, and both graduation and retention Odds Ratios and 95% Wald CIs are reported in Tables 4A-4I.

University	Likelihood Ratio χ^2 (p-value)	Coefficient of Determination R-Square
A	103.63 (<.0001)	0.1367
B	95.32 (<.0001)	0.0856
C	362.95 (<.0001)	0.1331
D	23.24 (.0007)	0.0004
E	842.31 (<.0001)	0.0713
F	237.21 (<.0001)	0.0278
G	111.73 (<.0001)	0.0623
H	200.58 (<.0001)	0.0280
I	19.57 (.0207)	0.0355

Table 2.G. *Graduation Models. Testing Global Null Hypothesis: β =0: Likelihood Ratio Chi-Squared Statistics (χ^2) and P-value (P), and Coefficient of Determination (R-Square).*

How well does the formulated multiple logistic regression model, as a whole, account for the dependent variables' behavior? This question is investigated by a likelihood ratio test for global null hypothesis: graduation likelihood does not depend on any of the six independent variables (β =0). The test yields a likelihood ratio chi-square statistic for each individual university's model. The likelihood ratio chi-square statistic, which is the analog of the F-statistic in a linear regression model, along with the p-value is reported in Tables 2.G and 2.R. The fact that all the p values for the likelihood ratio χ^2 are much smaller than 0.05 provides strong evidence against the global null hypothesis, indicating that the independent variables collectively predict graduation and retention at the 0.05 significant level.

University	Likelihood Ratio χ^2 (p-value)	Coefficient of Determination R-Square
A	177.46 (<.0001)	0.1620
B	154.33 (<.0001)	0.0977
C	732.94 (<.0001)	0.1620
D	647.74 (<.0001)	0.0820
E	1106.95 (<.0001)	0.0708
F	372.94 (<.0001)	0.0310
G	126.87 (<.0001)	0.0473
H	1195.40 (<.0001)	.0883
I	12.76 (.0004)	.0148

Table 2.R. Retention Models. Testing Global Null Hypothesis: β =0: Likelihood Ratio Chi-Squared Statistics (χ^2) and P-value (P), and Coefficient of Determination (R-Square).

The predictive efficacy of the model is examined by looking at the coefficient of determination, R-Square. R-Square represents the amount of variance in the dependent variable explained by the independent variables. For example, an R-Square value of 0.1367 for University A's graduation model indicates that 13.67% of the variance in graduation likelihood is due to the independent variables in the model: HSGPA and SATM. The R-Squares of graduation models and retention models are also included in Tables 2.G and 2.R, respectively.

Analysis and Results

Chi-Squared test statistics on the effects of the variables are reported in Tables 3.G and 3.R along with the p-values. For each variable, our question is two-fold: Does the given variable affect retention and/or graduation? And if so, does it predict retention differently than graduation?

For University A, HSGPA and SATM predict both graduation and retention with p-values generally less than 0.0001, which means the probability that HSGPA and SATM do not predict graduation and retention is less than 0.01%. For University B, while HSGPA and SATM predict graduation and retention, SATV, ETHNIC and CITIZEN predict retention. For University C, besides HSGPA and SATM, GENDER and SATV are found to be effective predictors of both graduation and retention, while ETHNIC and CITIZEN predict retention. For University D, HSGPA, SATM and SATV are not included in the model due to missing data. Among the three categorical data, ETHNIC is found to be significant to graduation and retention, while GENDER

University	GENDER χ^2 (p-val)	HSGPA χ^2 (p-val)	SATM χ^2 (p-val)	SATV χ^2 (p-val)	ETHNIC χ^2 (p-val)	CITIZEN χ^2 (p-val)
A	2.27 (0.13)	**42.83** (<0.0001)**	**18.87** (<0.0001)**	2.91 (0.08)	3.34 (0.50)	0.05 (0.81)
B	1.54 (0.21)	**51.00** (<0.0001)**	**11.28** (0.0008)**	1.69 (0.19)	6.53 (0.25)	0.06 (0.79)
C	**21.56** (<0.0001)**	**123.6** (<0.0001)**	**55.96** (<0.0001)**	**4.08** (0.04)**	5.60 (0.34)	2.13 (0.14)
D	1.49 (0.22)	Not tested	Not tested	Not tested	**19.06** (0.004)**	0.24 (0.62)
E	0.31 (0.57)	**464.94** (<0.0001)**	**171.61** (<0.0001)**	**25.21** (<0.0001)**	**25.35** (<0.0001)**	**31.15** (<0.0001)**
F	0.50 (0.47)	Not tested	**113.67** (<0.0001)**	**9.06** (0.002)**	**62.12** (<0.0001)**	0.07 (0.78)
G	**39.78** (<0.0001)**	**14.74** (0.0001)**	**19.99** (<0.0001)**	3.63 (0.05)	**16.83** (0.002)**	2.00 (0.36)
H	**6.99** (0.008)**	**70.35** (<0.0001)**	**46.82** (<0.0001)**	**35.77** (<0.0001)**	**13.65** (0.0085)**	**10.03** (0.0015)**
I	1.72 (0.18)	Not tested	**9.61** (0.0019)**	**6.88* (0.0087)**	4.70 (0.45)	1.62 (0.20)

Table 3.G. Graduation Analysis: *Type III Analysis of Effects: Wald Chi-Squared Statistics (χ^2) and P-value (P)*

University	GENDER χ^2 (p-val)	HSGPA χ^2 (p-val)	SATM χ^2 (p-val)	SATV χ^2 (p-val)	ETHNIC χ^2 (p-val)	CITIZEN χ^2 (p-val)
A	0.05 (0.81)	**61.71** (<0.0001)**	**14.66** (<0.0001)**	3.75 (0.05)	1.11 (0.89)	1.00 (0.31)
B	**4.82* (0.02)**	**26.61** (<0.0001)**	**10.44** (0.0014)**	**11.74* (0.006)**	**50.61* (<.0001)**	**11.06* (0.004)**
C	**29.08** (<0.0001)**	**203.99** (<0.0001)**	**78.50** (<0.0001)**	**9.77** (0.018)**	**49.74* (<.0001)**	**10.39* (0.005)**
D	**5.81* (0.01)**	Not tested	Not tested	Not tested	**42.31** (<.0001)**	**515.43* (<.0001)**
E	0.15 (0.69)	**607.99** (<0.0001)**	**217.98** (<0.0001)**	**5.739** (0.016)**	**38.63** (<0.0001)**	**35.68** (<0.0001)**
F	1.079 (0.29)	Not tested	**183.62** (<0.0001)**	**36.10** (<0.0001)**	**20.27** (0.0011)**	**5.608* (0.017)**
G	**22.09** (<0.0001)**	**21.82** (0.0001)**	**9.36** (0.002)**	0.133 (0.715)	**24.83** (0.0001)**	3.497 (0.174)
H	**27.00** (<0.001)**	**533.22** (<0.0001)**	**5.40** (0.02)**	**95.92** (<0.0001)**	**36.43** (<0.0001)**	**6.76** (0.009)**
I	0.29 (0.58)	Not tested	**4.97** (0.02)**	0.33 (0.56)	8.42 (0.13)	0.83 (0.36)

Table 3.R. Retention Analysis: *Type III Analysis of Effects: Wald Chi-Squared Statistics (χ^2) and P-value (P).*

and CITIZEN predict retention. For University E, all variables except GENDER are found to predict both graduation and retention. HSGPA is excluded in the study of University F, and among the five remaining variables, SATM, SATV and ETHNIC predict graduation and retention while CITIZEN predicts retention. GENDER, HSGPA, SATM and ETHNIC predict both graduation and retention for University G. In University H, all six variables are significant predictors of graduation and retention. Finally, for University I, SATM predicts graduation and retention, while SATV predicts graduation.

While all variables are found to be significant predictors of both graduation and retention for certain institutions, HSGPA and SATM are found to be predictors of both graduation and retention across all universities in which this data were available. It is interesting that GENDER, SATV, ETHNIC and CITIZEN are frequently predictors of retention but not graduation. This indicates a possible drawback to studies where the definition of retention includes students currently enrolled in an engineering program, in that variables that appear to be significant in the short-run (i.e., before graduation) may not, in fact, be significant in the final analysis.
The effects of the predictors are further quantified by the estimated β parameter values in the multiple logistic regression model. Odds Ratios are transformations of the β parameter estimates, which provide an intuitive view of how much students' odds of graduation and retention differ due to differences in the predictors. Tables 3A – 3I show Odds Ratio estimates and the Wald 95% confidence intervals for the Odds Ratio for the significant predictors listed in Tables 2.G and 2.R, with blank cells indicating no statistical significance. In what follows, we focus on the Odds Ratios for graduation.

For all Universities except D (and of course F and I, where HSGPA was not included), a marked Odds Ratio is associated with HSGPA, ranging from 1.279 to 3.83. This indicates that a 1-point increase in high school GPA increases likelihood of graduation by a factor of 1.27 to 3.83. GENDER was significant to graduation (although not consistently positively or negatively) for Universities C, G and H, where the Odds Ratios for female vs. male were 0.53, 1.925 and 0.853, respectively. This means that a female's likelihood of graduation in University G is nearly twice that of a male's, while in Universities C and H, a female's likelihood of graduation is below that of a male's, 0.53 and 0.853, respectively. For all universities that included SATM, the Odds Ratios varied from 1.003 to 1.006, suggesting that, not surprisingly, math SAT scores correlate positively with graduation. Specifically, a 100-point increase in math SAT score results in a 30 to 60 percent increase in likelihood of graduation. Interestingly, the Odds Ratios for SATV varied from 0.997 to 0.999, indicating that verbal SAT scores correlate negatively

Parameter	Odds Ratio For Graduation	95% CI on Odds Ratio	Odds Ratio For Retention	95% CI on Odds Ratio
HSGPA	3.63	2.46 – 5.34	3.19	2.38-4.26
SATM	1.005	1.003 – 1.008	1.004	1.002-1.006

Table 4A. *Odds Ratio estimates and 95% Wald confidence intervals for graduation and retention: Significant predictors for University A.*

with retention. Ethnicity played a role in graduation in six of the Universities, but with the ethnic group having highest likelihood of graduation strongly dependent on the institution. Finally, citizenship was a significant predictor for Universities E and H. In University E, the Odds Ratio of NRAlien vs. ResAlien was 2.343, while in University H, the Odds Ratio of Citizen vs. NRAlien was 2.722.

Parameter	Odds Ratio For Graduation	95% CI on Odds Ratio	Odds Ratio For Retention	95% CI on Odds Ratio
GENDER (Female vs. Male)			0.721	0.539-0.965
HSGPA	3.83	2.649 – 5.537	1.81	1.446-2.271
SATM	1.004	1.002 – 1.006	1.003	1.001-1.005
SATV			1.003	1.001-1.004
ETHNIC (AfrAm vs. White)	1.711	1.089 – 2.687	3.295	2.336-4.591
CITIZEN (Citizen vs. ResAlien)			0.08	0.17-0.367
CITIZEN (NRAlien vs. ResAlien)			0.066	0.013-0.336

Table 4B. *Odds Ratio estimates and 95% Wald confidence intervals for graduation and retention: Significant predictors for University B.*

Parameter	Odds Ratio For Graduation	95% CI on Odds Ratio	Odds Ratio For Retention	95% CI on Odds Ratio
GENDER (Female vs. Male)	0.53	0.405 – 0.693	0.613	0.513 – 0.732
HSGPA	3.491	2.8 – 4.351	3.232	2.751 – 3.796
SATM	1.006	1.004 – 1.007	1.005	1.004 – 1.006
SATV	0.999	0.997 – 1.0	1.001	1.001 – 1.002
ETHINIC (AfrAm. vs. White)			2.34	1.765 – 3.104
ETHINIC (Asian vs. White)			1.329	1.015 – 1.74
ETHNIC (Hisp vs. White)			1.751	1.378 – 2.225
CITIZEN (Citizen vs. ResAlien)			0.453	0.258 – 0.795

Table 4C. *Odds Ratio estimates and 95% Wald confidence intervals for graduation and retention: Significant predictors for University C.*

Parameter	Odds Ratio For Graduation	95% CI on Odds Ratio	Odds Ratio For Retention	95% CI on Odds Ratio
GENDER (Female vs. Male)			0.865	0.769 – 0.973
ETHNIC (AfrAM vs. White)	0.671	0.559 – 0.806		
ETHNIC (NRAlien vs. White)			0.257	0.163 – 0.404
CITIZEN (ResAlien vs. Citizen)			5.651	4.866 – 6.562

Table 4D. *Odds Ratio estimates and 95% Wald confidence intervals for graduation and retention: Significant predictors for University D.*

Parameter	Odds Ratio For Graduation	95% CI on Odds Ratio	Odds Ratio For Retention	95% CI on Odds Ratio
HSGPA	3.426	3.064 – 3.832	3.523	3.194 – 3.886
SATM	1.004	1.004 – 1.005	1.004	1.004 – 1.005
SATV	0.999	0.998 – 0.999	0.999	0.999 – 1.000
ETHNIC (AfrAm. vs. White)			1.177	1.035 – 1.340
ETHNIC (Asian vs. White)			1.190	1.043 – 1.357
ETHNIC (Hisp. vs. White)	1.706	1.37-2.21	1.729	1.426 – 2.096
CITIZEN (NRAlien vs. ResAlien)	2.343	1.599-3.432	1.834	1.342 – 2.508

Table 4E. *Odds Ratio estimates and 95% Wald confidence intervals for graduation and retention: Significant predictors for University E.*

Parameter	Odds Ratio for Graduation	95% CI on Odds Ratio	Odds Ratio For Retention	95% CI on Odds Ratio
SATM	1.004	1.003 – 1.005	1.005	1.004 – 1.006
SATV	0.999	0.998 – 1.000	1.002	1.001 – 1.002
ETHNIC (AfrAM vs. White)	0.417	0.329 – 0.530	0.710	0.583 – 0.863
ETHNIC (Intrnat vs. White)	1.801	1.079 – 3.004		
CITIZEN (Citizen vs. NRAlien)			0.724	0.555 – 0.944

Table 4F. *Odds Ratio estimates and 95% Wald confidence intervals for graduation and retention: Significant predictors for University F.*

Parameter	Odds Ratio For Graduation	95% CI on Odds Ratio	Odds Ratio For Retention	95% CI on Odds Ratio
GENDER (Female vs. Male)	1.925	1.571 – 2.360	1.487	1.260 – 1.754
HSGPA	1.279	1.128 – 1.450	1.371	1.201 – 1.564
SATM	1.003	1.002 – 1.005	1.002	1.001 – 1.003
ETHNIC (AfrAM vs. White)	2.618	1.594 – 4.299	2.555	1.756 – 3.716
ETHINIC (Asian vs. White)	8.347	1.409 – 49.46		

Table 4G. *Odds Ratio estimates and 95% Wald confidence intervals for graduation and retention: Significant predictors for University G.*

Parameter	Odds Ratio for Graduation	95% CI on Odds Ratio	Odds Ratio For Retention	95% CI on Odds Ratio
GENDER (Female vs. Male)	0.853	0.759 – 0.960	0.790	0.723 – 0.864
HSGPA	1.758	1.541 – 2.006	3.049	2.774 – 3.352
SATM	1.003	1.002 – 1.003	1.001	1.000 – 1.001
SATV	0.998	0.997 – 0.999	1.003	1.002 – 1.003
ETHNIC (AfrAM vs. White)	0.756	0.632 – 0.906	1.308	1.146 – 1.492
ETHNIC (Asian vs. White)			1.362	1.142 – 1.626
ETHNIC (Hisp. vs. White)			1.696	1.201 – 2.397
ETHNIC (NatAm. vs. White)			1.576	1.038 – 2.393
CITIZEN (Citizen vs. NRAlien)	2.722	1.465 – 5.059	1.886	1.169 – 3.041

Table 4H. *Odds Ratio estimates and 95% Wald confidence intervals for graduation and retention: Significant predictors for University H.*

Parameter	Odds Ratio for Graduation	95% CI on Odds Ratio	Odds Ratio For Retention	95% CI on Odds Ratio
SATM	1.005	1.002 – 1.008	1.003	1.000 – 1.005
SATV	0.997	0.994 – 0.999		

Table 4I. *Odds Ratio estimates and 95% Wald confidence intervals for graduation and retention: Significant predictors for University I*

Conclusion

We find that graduation in engineering for students who enter in an engineering discipline depends significantly upon several factors. High school GPA and math SAT scores were positively correlated with graduation rates for all universities for which this data were available. Interestingly, verbal SAT scores correlated negatively with odds of graduation for five out of eight universities. While gender, ethnicity and citizenship also showed significant effects these were not consistently positive or negative. In two universities, the graduation rate for males was higher than that for females, while in one university the graduation rate was higher for females. Ethnicity was significant in six universities. Finally, in two of the universities citizenship significantly affected graduation.

By including all students *enrolled* in the last year of the study, in addition to those having already graduated, we are able to contrast so-called retention with graduation. Retention is also significantly influenced by gender, high school GPA, math SAT scores, verbal SAT scores, ethnicity and citizenship. However, we find that gender, verbal SAT, ethnicity and citizenship frequently appear as significant predictors of retention where they do not appear significant to graduation. This suggests one must be careful in defining success in these longitudinal studies, as variables which appear significant in the short run (i.e., before graduation) may not, in fact, be significant in the longer run.

References

[1,3] Astin, A.W. (1971) *Predicting Academic Performance in College,* The Free Press, New York.
Besterfield-Sacre, M., Moreno, M., Shuman, L.J. and Atman, C.J. (2001) "Gender and Ethnicity Differences in Freshmen Engineering Student Attitudes: A cross-Institutional Study," Journal of Engineering Education October pp 477-489.

[2] Seymour, E., and Hewitt, N. M. (2000) *Talking About Leaving: Why Undergraduates Leave the Sciences*, Westview Press.

[4] Zhang, Z. and RiCharde, R.S., (1998) "Prediction and Analysis of Freshman Retention" *AIR 1998 Annual Forum Paper* Minneapolis, MN.

[5] Levin, J. and Wyckoff, J. (1990) "Identification of Student Characteristics that Predict Persistence and Success in an Engineering College at the End of the Sophomore Year: Informing the Practice of Academic Advising," *Division of Undergraduate Studies Report No. 1990.1* Pennsylvania State University.

[6] LeBold, W. K. and Ward, S. K. (1998) "Engineering Retention: National and Institutional Perspectives," *Proceedings, 1988 ASEE Annual Conference*. ASEE, 1988. pp 843-851.

[7] Agresti, A. (1996) *An Introduction to Categorical Data Analysis*, John Wiley & Sons, Inc., New York.

Author Information

Guili Zhang is a Ph.D. candidate in Educational Research and Statistics, Department of Educational Psychology, University of Florida. She received a B.A. in British and American Language and Literature at Shandong University, China, and a M.Ed. in English Education at Georgia Southern University. She has published extensively and has won numerous awards at the national and regional level in the area of educational research in China. She teaches Measurement and Assessment in Education at the University of Florida. Her research interests involve applied quantitative research designs, categorical data analysis, and structural equation modeling.

Tim Anderson is Chairman and Professor in the Department of Chemical Engineering, University of Florida. He is also Director of the SUCCEED Engineering Education Coalition and editor of *Chemical Engineering Education*. He received a Ph.D. at the University of California-Berkeley in 1979. His discipline research interests are in the area of electronic materials processing.

Matthew Ohland is Assistant Professor in General Engineering at Clemson University. He has a Ph.D. in Civil Engineering with a minor in Education from the University of Florida, and served as the Assistant Director of the SUCCEED Coalition until 2000. His research is in freshman programs and educational assessment. He has been elected to the Executive Council of Tau Beta Pi, the engineering honor society, and will be installed as its President in October 2002.

Rufus Carter is a Ph.D. student in Educational Research and Testing, Department of Educational Psychology, University of Florida. He received a B.S. in Psychology and Sociology from the University of Virginia, Wise. His research interests involve test and survey validation, generalizability of high stakes performance exams, classroom and project assessment and evaluation, and methodologies for setting performance standards.

Brian Thorndyke is a Ph.D. candidate in the Quantum Theory Institute, Department of Physics, University of Florida. He received a M.Sc. in High Energy Physics at the University of Montreal, and an M.S. in Computer Science at the University of Florida. His research interests involve computational methods applied to the dynamics of mixed quantum/classical molecular systems.