# AC 2012-4377: MODELING STUDENT SUCCESS OF INTERNATIONAL UNDERGRADUATE ENGINEERS

**Dr. Teri Reed-Rhoads, Purdue University, West Lafayette**

Teri Reed-Rhoads is Assistant Dean of engineering for undergraduate education, Associate Professor in the School of Engineering Education, and Director of the First-year Engineering program at Purdue University. She received her B.S. in petroleum engineering from the University of Oklahoma and spent seven years in the petroleum industry, during which time she earned her M.B.A. She subsequently received her Ph.D. in industrial engineering from Arizona State University. Reed-Rhoads' teaching interests include statistics, interdisciplinary and introductory engineering, diversity, and leadership. Her research interests include statistics education, concept inventory development, assessment/evaluation of learning and programs, recruitment and retention, diversity, equity, and cultural humility. She has received funding from the National Science Foundation, Department of Education, various foundations, and industry. Reed-Rhoads is a member and Fellow of the American Society for Engineering Education and a member of the Institute of Electronics and Electrical Engineers and the Institute of Industrial Engineers. She serves as an ABET EAC Evaluator for ASEE.

**Dr. P.K. Imbrie, Purdue University, West Lafayette**

P.K. Imbrie is an Associate Professor of engineering in the Department of Engineering Education at Purdue University. He holds B.S., M.S., and Ph.D. degrees in aerospace engineering from Texas A&M University. His research interests include educational research, solid mechanics, experimental mechanics, microstructural evaluation of materials, and experiment and instrument design. He has been involved with various research projects sponsored by NSF, NASA, and AFOSR, ranging from education-related issues to traditional research topics in the areas of elevated temperature constitutive modeling of monolithic super alloys and environmental effects on titanium based metal matrix composites. His current research interests include epistemologies, assessment, and modeling of student learning, student success, student team effectiveness, and global competencies; experimental mechanics; and piezospectroscopic techniques.

**Qu Jin, Purdue University, West Lafayette**

Qu Jin is a graduate student in the School of Engineering Education at Purdue University. She received a M.S. degree in biomedical engineering from Purdue University and a B.S. degree in material science and engineering from Tsinghua University in China. Her research focuses on modeling student success outcomes, which include placement, retention, academic performance, and graduation.

**Mr. Joe J.J. Lin, Purdue University, West Lafayette**

## Modeling Student Success of International Undergraduate Engineers

Modeling student retention using entering secondary school academic performance metrics only is limited at best. Past research has shown that these variables can be somewhat informative, but are not the whole story. In order to expand our understanding of successful students, defined in this study as students who are retained and ultimately graduate with a degree in engineering, student retention and graduation modeling has been extended to include not only secondary school academic performance, but also self-reported affective and attitudinal measures. The Student Attitudinal Success Instrument (SASI), a 161-item survey assessing 13 specific noncognitive constructs, was developed based largely on existing instruments. The SASI is designed to provide data on noncognitive characteristics for incoming engineering students (a) *prior to* the onset of the first year and (b) for which higher education institutions may have an influence during students' first year. Data collected from this instrument have been found to be suitable for use in the development of predictive models of student retention and/or graduation, which is the definition of success in this model. The SASI is used to provide information about the academic preparation and affective characteristics of incoming first-year engineering students. Such systematically gathered information helps us assess the impact of University and programmatic decisions aimed at student recruitment, admission, retention, and ultimately the success of all students and, in particular, minority student populations.

Though international students in engineering tend to have higher levels of overall retention and graduation versus any other majority or minority population, this study shows that the trend is in a concerning downward direction. In order to reverse this graduation trend, programs need to be expanded or created that are based on informed data decisions specific to student populations, such as international students. Understanding additional measures beyond admission metrics that lead to student success allows policy, programs or programmatic changes that increase overall student success. This study begins with a review of how this type of modeling was used to inform a change in admissions policy in the case of gender bias. Then, the techniques are expanded to international student success modeling. Though international and domestic students report similar levels of each success measure, the relative importance of each measure in predicting retention was different for these two student populations.

### Introduction

Retention and graduation rates have become increasingly important as calls for changes in higher education increase. At the same time, the number of international students enrolled in institutions of higher education in the United States has increased from 547,867 in the 2000/01 academic year to 723,277 in the 2010/11 academic year[1]. Engineering remains the second largest field of study for international students behind Business and Management[1]. The number of international students in engineering during the 2008/9 academic year was 118,980 and increased to 135,592 in the 2010/11 academic year[1,2]. In general, international students are noted to achieve higher levels of retention and ultimately graduation as compared to domestic students[3]. However, the tracking of graduation trends at a major Midwest public institution reveals several concerning trends. These include falling 4-year (Figure 1) and 6-year (Figure 2) graduation rates. The 4-year graduation rate has fallen from 70% to just over 40% in the past 15 years. Another trend was confirmed and that is that international students at this institution who start in engineering either finish in engineering or leave the university. This can be seen from the lack of gap between the engineering and university graduation trends in both Figures 1 and 2. On the positive side, international student graduation rates, at the 6-year time frame, are still significantly higher than the domestic student graduation rates. However, the falling trends in graduation rates coupled with an increase in number of international students (as a percentage of the overall student population) implies we can no longer take for granted international student success and research needs to be reported on retention mechanisms directly impacting our international student populations. Based on this, we present an initial examination of the factors that could be addressed through various programming efforts. Prior to the international student modelling results, we present an example of a positive outcome from the use of modelling related to another minority population; women.

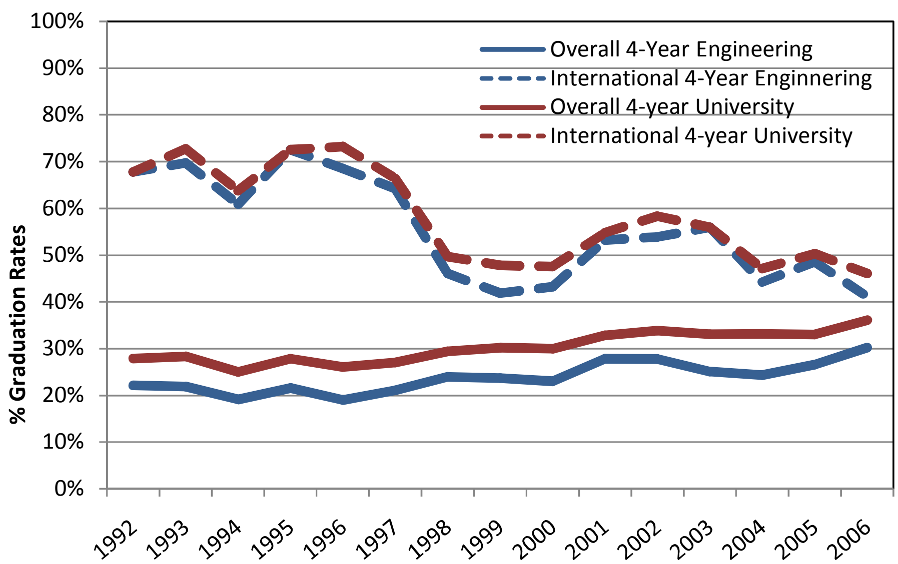


Figure 1. Comparison of 4-year engineering and university graduation rates of domestic cohorts and international students

A discussion of academic admissions practices quickly gets to the question of what are indicators of student success. The difficulty of this discussion is that historically the factors used to answer this question have been cognitive, such as high school grade point averages, class rank, and standardized test scores. However, research indicates that there are many factors affecting retention. For example, Astin[4] showed that students who tend to be more engaged are more likely to persist; further, well over half of the variation between institutions on a measure of "student engagement" can be attributed to characteristics present *prior to* entering college. Fostering a spirit of engagement in students who may not show that propensity may increase students' chances for success, and intervention program(s) designed to increase student engagement are certainly feasible from an institution during the first year of study. It should be noted that the majority of Astin's work is based on domestic students, thus there exists a limitation of applying this framework potentially to the international student population. Perhaps more importantly, the mention of modelling to discuss admissions policies elicits strong concern that profiling will result and will limit admissions to certain populations. Weinstein, et.

al,[5] acknowledge this modeling controversy and relate it to being a misconception that "the role of models is to establish truth rather than to guide clinical and policy decisions" (pg. 348). These

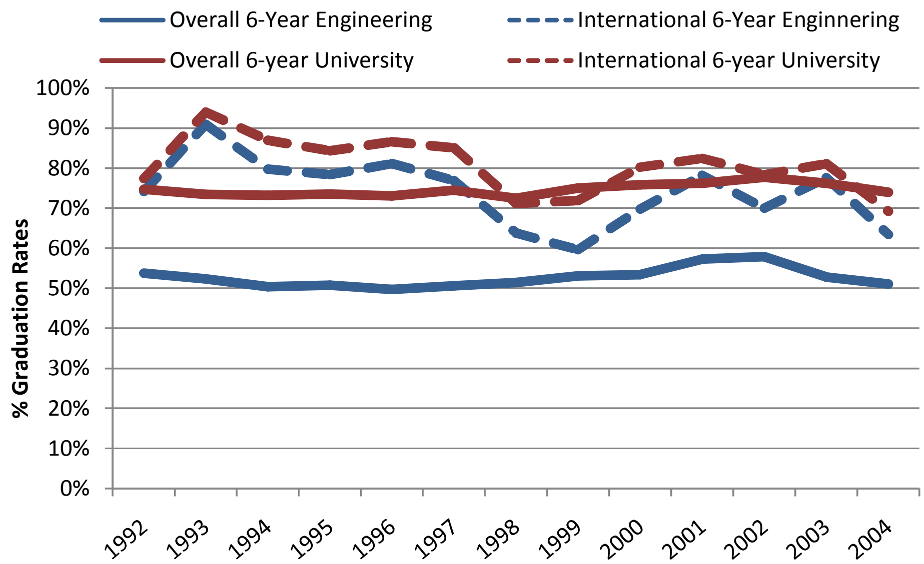


Figure 2. Comparison of 6-year engineering and university graduation rates of domestic cohorts and international students

authors provide examples of public policy domains from areas such as environmental protection or defense strategy that involve human life and health where models are generally accepted as decision aids. We propose the same use of modeling in student retention to view admission policies and academic support programming in an advisory capacity.

**Modeling Framework**

The SASI[6-8], a 161-item survey assessing 13 specific noncognitive constructs was developed based largely on existing instruments. This instrument is designed to provide data on noncognitive characteristics for incoming engineering students (a) *prior to* the onset of the first year and (b) for which higher education institutions may have an influence during students' first year. Data collected from this instrument may be suitable for use in the development of predictive models of student retention and/or graduation, which is the definition of success in this model. In this study, modeling is completed for first year retention and then graduation after 4 years (8 semesters) and 5 years (10 semesters). Due to such factors as the number of credit hours to obtain an engineering degree or a student's participation in cooperative learning (one semester in school alternated with one semester in industry), average graduation rates in the US are just over 4 years thus the 5 year graduation interest.

The SASI is used to provide the College information about the academic preparation and affective characteristics of incoming first-year engineering students. Such systematically gathered information helps us assess the impact of University and programmatic decisions aimed at student recruitment, admission, retention, and ultimately the success of all students and, in particular, underrepresented students. Underrepresented students in the US in the field of engineering includes women, as well as race and ethnicity designations of African American, Hispanic/Latino American, Native American, Native Hawaiian and Pacific Islander for US domestic students only.

The current model of student success includes 13 self-reported affective factors (leadership, deep- vs. surface-learning types, team or individual orientation, academic self-efficacy, motivation, meta cognition, expectancy value, major decision, goal orientation, implicit beliefs, intent to persist, social climate, and self worth); and 11 academic preparation items from high school including overall GPA, core GPA, standardized test results by subarea [SAT/ACT], average grades in mathematics, science, and English, and the number of semesters completed of mathematics, science, and English. Table 1 provides details of the origins of each factor.

Table 1. References used during the development of the SASI.

| Factor | Reference(s) |
|---|---|
| Motivation | Pintrich and Schunk[9], French and Oakes[10] |
| Metacognition | Pintrich and DeGroot[11], O'Neil and Abedi[12] |
| Deep learning | Biggs, Kember and Leung[13] |
| Surface learning | Biggs, Kember and Leung[13] |
| Academic Self-efficacy* | Bandura[14], Pajares[15] |
| Leadership* | Hayden and Holloway[16] |
| Team vs. Individual Orientation* | McMaster[17] |
| Expectancy-Value | Wigfield and Eccles[18] |
| Major Indecision[+] | Osipow[19] |

* Developed internally based upon the cited reference(s).
[+] Originally developed and presented as a Career Indecision scale by the Osipow[19] and modified to be an Engineering Major Indecision scale.

Figure 3 provides an example of prior research wherein modeling was used to identify important factors for predicting “success” of male and female, where success was operationalized as “1 year retention,” “8 semester graduation,” or “10 semester graduation, respectively. The factor data (the independent variables) were collected just prior to students beginning their college experience. The importance of a particular factor towards predicting the outcome variable (success) is indicated by the radial distance from the center of the circle (center - low importance, perimeter of the circle – high importance). The “1 Year Retention” radial plot in Fig. 3, indicates leadership is an important attribute for women’s success (defined as retention) at the first-year level. For men, the semesters of high school mathematics is important. Semesters of science is a factor important to both men’s and women’s success. Since the cohort remains the same for all 3 graphs, the results indicate that factors important to predict success for women and men are not the same and that important factors change with the measure of success. Since the model is used for predictive purposes rather than an explanatory tool, the authors can offer no illuminating rationale as to why the important factors differ by sex.

As a result of legal rulings in the United States, admissions policies can use a holistic mix of factors in directing admissions and scholarship decisions. However, different factors cannot be used for men and women, or majority and minority students, even if these factors were to be based on the success factor data for each population. The authors have used the success factor data in developing our admissions factor recommendations to future policies that place a higher emphasis on affective indicators.

Thus, based on this model and additional research done on admission factors, a policy change recommendation was made that admission decisions be made using a set of priorities that include cognitive factors such as SAT verbal and number of semesters of mathematics, science, and English while taking into account leadership, major decision, and academic motivation. In addition, based on reports such as the National Academy of Engineering's report, *Changing the Conversation*[20], social relevance of engineering as a discipline has been added to the decision process.

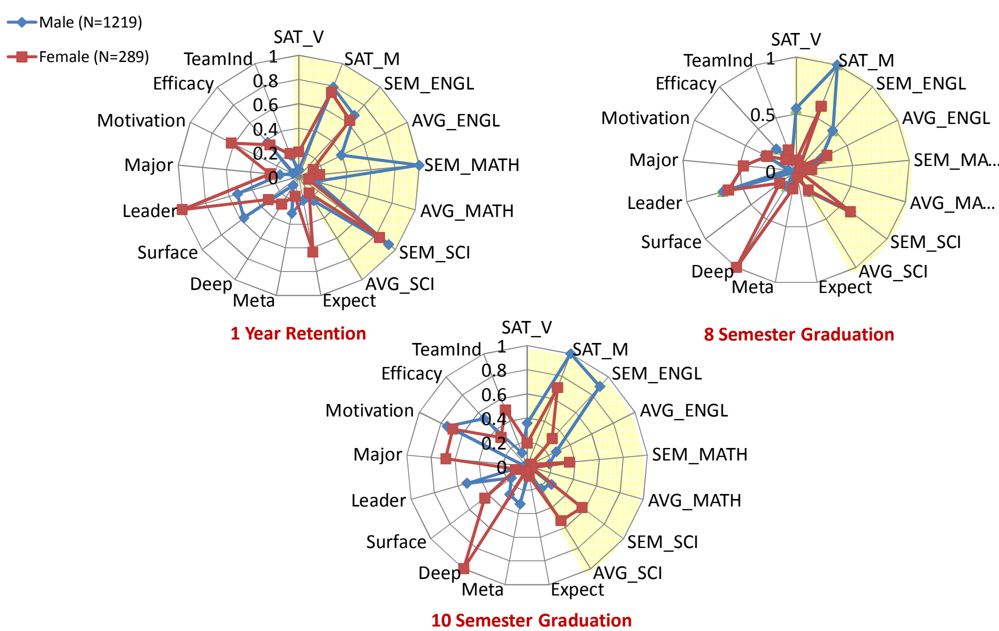


Figure 3, SASI model of success factors by sex for the 2004 cohort.

## Research Methods

This section will discuss the population of participants, the instruments employed, independent and dependent variables, and proposed modeling and analysis methodologies.

### *Participants*

The participants in this study included 3,001 first-year engineering students from 2004 and 2005 cohorts at a large Midwestern university. These particular cohorts were chosen since they are the most recent that can be used to obtain 8/10/12 semester graduation along with first-year retention. Among them, 17.7% are females and 82.3% are males. Ethnicity was as follows: 81.0% Caucasian, 7.8% Asian/Pacific Islander, 2.2% African American, 1.3% Hispanic, 0.5% American Native, 5.3 % International, and 1.8% others. Overall, there are 158 international students (5.3%) and 2843 U.S. domestic students (94.7%).

### *Instrumentation*

The Student Attitudinal Success Instrument (SASI)[6-8] is applied in this study just as it was in the gender research previously discussed. Recall that it is a Likert-style survey developed by researchers from a Midwestern university to collect self-reported information on student's various attitudinal and affective self-beliefs[6-8]. The first phase of SASI focuses on the following nine constructs: Leadership, Deep Learning, Surface Learning, Teamwork, Academic Self-efficacy, Motivation, Meta-cognition, Expectancy-value, and Major Decision. In 2007, SASI was expanded to fourteen constructs by adding five new factors: Goal Orientation, Implicit Beliefs, Intent to Persist, Social Climate and Self Worth. In this research, we focus our analysis on students from 2004 and 2005 cohorts, therefore the data are collected from the original version of SASI with nine original constructs.

In the original SASI, all Cronbach's coefficient alphas for these scales (constructs) were greater or equal to 0.80, except for the Teamwork scale ( alpha =.74). Multiple studies have supported the scales' construct validity based on the results of confirmatory factor analyses[6-8].

### *Independent Variables*

The independent (predictor) variables collected in this study include: 1) eight items from student's high school performance measures, and 2) eight affective and attitudinal self-belief constructs from SASI survey. The high school performance measures include: standardized test results (verbal and math), average high school grades in mathematics, science, and English classes, and also the number of semesters in mathematics, science, and English in high school. The eight attitudinal and affective self-beliefs applied include Leadership, Deep Learning, Surface Learning, Teamwork, Self-efficacy, Meta-cognition, Expectancy-value, and Major decision. The construct Motivation from original SASI was not used in this study, due to a very high correlation (0.80) with Self-efficacy in this data set.

### *Dependent Variables*

There are four dependent variables in this study. They are: 1) student's retention in engineering after one year, 2) 8-semester graduation in engineering, 3) 10-semester graduation in engineering, and 4) 12-semester graduation in engineering.

The authors have collected students' enrollment statuses at the beginning of every semester following their freshman year. For measuring retention, students remaining in the lower-division and upper divisions (specific disciplines) engineering programs after first year were considered as "retained" students. The students who transferred to majors other than engineering, or left the university completely were classified as "not-retained". For 8/10/12-semester graduation status, the researchers collected each student's enrollment status after 8/10/12 semesters and determine if they successfully completed their bachelor degrees in engineering after these numbers of semesters.

### *Method for Modeling Student Retention and Graduation*

Logistic regression is a statistical method frequently used in educational studies to predict student's retention or graduation status. It is a modeling tool mostly used for data analysis and inference, when the goal is to understand the role of input variables in explaining the outcome[21]. When the outcome (dependent variable) is binary, such as the students' retention or graduation status, traditional multiple regression is not appropriate as the modeling tool. Instead, logistic regression becomes a more proper statistical method for such models[22, 23]. Discriminant analysis is another modeling tool capable of handling binary dependent variables. However, previous researchers have reported better performance and advantages from logistic regression when compared with discriminant analysis[24-25]. Another potential tool for modeling retention/graduation is neural networks. Some scholars reported neural networks models produce similar or better prediction/classification results in comparison with logistic regression. However, neural networks model requires large number of training samples to take advantage of its flexibility[26]. In this study with a special focus on international students, we have a limited number of foreign students in our data set. This inevitably become a constraint while choosing the modeling methods. For logistic regression, several researchers have suggested a minimum observation-to-predictor ratio of 10 to 1. This moderate data requirement allows us to use logistic regression in this study with limited international student population.

To develop a logistic regression model:

Assuming

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_n x_n + e, \tag{1}$$

Where

| | |
|---|---|
| $y$ | is the dependent variable |
| $x_i$ | is the $i^{th}$ independent variable |
| $\beta_0$ | is the regression constant (intercept) |
| $\beta_i$ | is the $i^{th}$ regression coefficient |
| $n$ | is the number of independent variables, |
| $e$ | is the error |

In general, a logistic regression model computes the class membership probability for one of the two possible outcomes in the dependent variable. A logistic regression model (with 1-year retention as dependent variable) can be formulated mathematically as below:

The probability of freshman student who stayed in engineering program after first year, E(X), can be expressed by sigmoidal function as:

$$E(X) = \frac{Exp(y')}{1 + Exp(y')} \text{, where } y' = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_n x_n \tag{2}$$

i.e.:

$$E(X) = \frac{Exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_n x_n)}{1 + Exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_n x_n)} \tag{3}$$

The sigmoidal function is attractive in this type of application because it shows asymptotes $E(-\infty) = 0$ and $E(\infty) = 1$, while approaches linearity in the middle range around $E(X) = 0.5$.

Coefficients ($\beta_i$) in the model are estimated by the maximum likelihood method. The purpose is to determine the coefficients $\beta_i$ so that $E(X)$ is the best fit to the data. A larger regression coefficient in logistic regression generally suggests the variable have a higher impact on the outcome (dependent variable) than other variable with a smaller coefficient.

### *Models for Different Outcomes and Student Population*

In this study, logistic regression models are developed to study the impact of independent variables on these four dependent variables (outcomes) we discussed earlier: 1) student's retention in engineering after one year, 2) 8-semester graduation in engineering, 3) 10-semester graduation in engineering, and 4) 12-semester graduation in engineering. Every model with specific outcome will be developed twice, once with international students and the other with U.S. domestic students. After the models were developed with student data, the regression coefficients are studied and compared to explain the relative importance of each independent variable on the outcome of the model.

## Results of International Student Modeling

A comparison of success factors for international versus domestic students for the 2004 and 2005 cohorts was performed (these cohorts were prior to a large increases in international student numbers at many institutions within the United States). As was found when comparing success factors by gender, the international student population was noted to have different success indicators from that of domestic students. As shown in Fig.4, the most significant factor found for predicting international student success, where success was operationalized as retained after the first year in the College of Engineering as well as 8/10/12 semester graduation, was expectancy value. This factor is based on the work of Wigfield and Eccles[18] and represents the influence of expectancies and values on student achievement. The scale consists of 32 items in five subscales: Expected Use of Academic Resources (5 items), Community Involvement (4 items), Employment Opportunities (8 items), Persistence (as related to academics – 7 items) and Social Engagement (8 items). In contrast, the most significant factor for domestic students was the standardized mathematics test score (SAT_M). Interestingly, SAT_M was the least significant factor for international students. Overall, the top five factors for international students included: expectancy-value, the average grade in mathematics and science courses during their secondary education, academic self-efficacy and the number of semesters of science during their secondary education, respectively. In contrast, the significant factors for domestic students included: SAT mathematics test score, followed by the number of semesters of math during their secondary education, academic self-efficacy, the number of semesters of science during their secondary education and metacognition, respectively. Unlike female students in Fig. 3, the principle important success factors for both international students and domestic students were in the cognitive dimension. However, by the time graduation is achieved, expectancy-value is still important for international students while the semesters of mathematics and leadership have taken over the next two important factors. It should also be noted that leadership has now converged for both international and domestic students as an important factor.

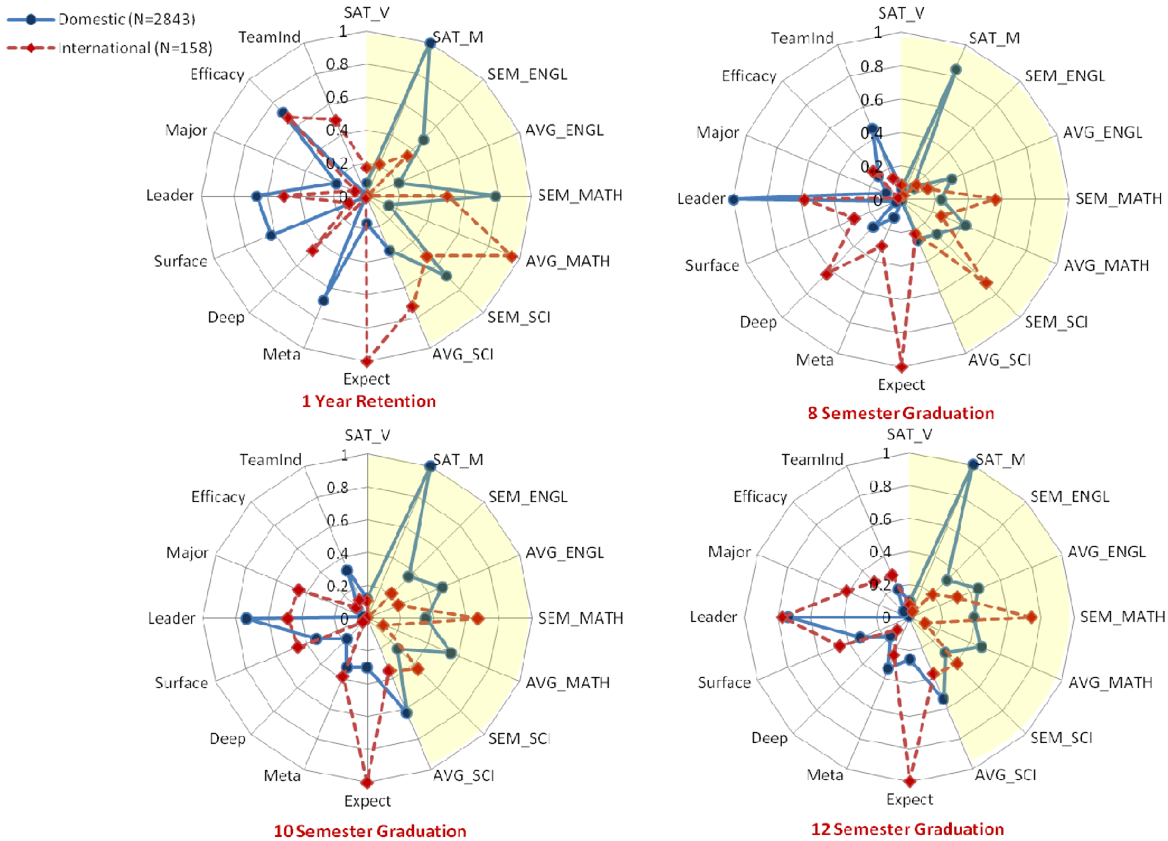


Figure 4. SASI model of success factors by residency for the aggregated 2004 and 2005 cohorts, ($N$=3001).

**Conclusions and Future Work**

The resulting struggle is now the validity of the application to provide indicators of some of these attributes. Because of the data-based nature and the breadth of this work, these results have been used to inform changes in admissions practices on the studied campus.  A more gender neutral set of admissions practices is essential for gender parity in engineering, particularly in light of recent research which suggests recruiting women into engineering is a larger issue than retaining women to graduation.  New international student retention programs and efforts can be informed with the results of this work.  For this research, all international students were pooled with no regard to region of the world.  Future work could investigate potential areas much like future work related to other under-represented populations can key on particular ethnicity or race.  Finally, as overall retention and graduation rates continue to fall and the number of international students enrolling continues to increase, further research of this under-represented population is warranted.

## References


1. Open Doors Fast Facts. *Open Door Report*, <http://www.iie.org/opendoors> (2011).
2. Open Doors Fast Facts. *Open Door Report*, <http://www.iie.org/opendoors> (2010).
3. Bista, K. & Foster, C. Issues of International Student Retention in American Higher Education. *The International Journal of Research and Review* **7**, 1 - 10 (2011).
4. Astin, A. Making Sense Out of Degree Completion Rates. *Journal of College Retention* **7**, 5 - 17 (2005).
5. Weinstein, M. C. *et al.*, (2001). Modeling for Health Care and Other Policy Decisions: Uses, Roles, and Validity. *Value in Health* **4**, 348 - 361.
6. Immekus, J., P. K. Imbrie and S. Maller, (2004). Work in progress: The influence of pre-college factors on first-year engineering students' academic success and persistence. *Proceedings of the Frontiers in Education Conference*, Savannah, GA.
7. Immekus, J., S. Maller, P. K. Imbrie, N. Wu, and P. McDermott, (2005). Work in progress: An analysis of students' academic success and persistence using pre-college factors. *Proceedings of the Frontiers in Education Conference*, Indianapolis, IN.
8. Reid, K. and P. K. Imbrie, (2008). Noncognitive characteristics of incoming engineering students compared to incoming engineering technology students: A preliminary examination. *Proceedings of the American Society for Engineering Education National Conference*, Pittsburgh, PA.
9. Pintrich, P. R., and Schunk, D. H. (1996). Motivation in education: *Theory, research, and applications*. Englewood Cliffs, NJ: Prentice Hall.
10. French, B.F., and Oakes, W. (2001). Measuring Academic Intrinsic Motivation: Evidence of Reliability and Validity of a New Instrument, presented at the Annual Conference for the American Education Research Association, Seattle, WA, 2001.
11. Pintrich, P. R., and De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82: 33-40.
12. O'Neil, H.F., and Abedi, J. (2000). The State Metacognitive Inventory. National Center for Research on Evaluation Standards, and Student Testing, Univ. California, LA, CA, Los Angeles CSE Tech Report, 469.
13. Biggs, J., Kember, D. and Leung, D.Y.P. (2001). The Revised Two-Factor Study Process Questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*, 71:133-149.
14. Bandura, A. (1986). Social Foundations of Thought and Action: *A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall.
15. Pajares, F. (1996). Self-Efficacy Beliefs in Academic Settings. *Review of Educational Research*. 66: 543-578.
16. Hayden, D. and Holloway, E. (1985). A Longitudinal Study of Attrition Among Engineering Students. *Journal of Engineering Education*, 71(7): 664-668.
17. McMaster, J., (1996). Desired attributes of an engineering graduate. *Proceedings of the AIAA Advanced Measurement and Ground Testing Technology Conference*, New Orleans, LA.
18. Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68–81.
19. Osipow, S. (1999). Assessing Career Indecision. *Journal of Vocational Behavior*, 55 (1): 147-154.
20. National Academy of Engineers, (2008). *Changing the Conversation: Messages for Improving Public Understanding of Engineering*, The national Academies Press, Washington, D.C.
21. Hastie T, Tibshirani R, Friedman J., (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed: Springer; 2009.
22. Schaeffers KG, Epperson DL, Nauta MM, (1997). Women's Career Development: Can Theoretically Derived Variables Predict Persistence in Engineering Majors? *Journal of Counseling Psychology*. 1997;V. 44:pp. 173-183.
23. Cody RP, Smith JK, (2006). *Applied statistics and the SAS programming language*. 5th ed. Upper Saddle River, N.J.: Pearson Prentice Hall; 2006.
24. Press SJ, Wilson S., (1978). Choosing Between Logistic Regression and Discriminant Analysis. *Journal of the American Statistical Association*. 1978;73(364):pp.699-705.
25. Hair JF, (1995). *Multivariate data analysis*, 4th ed. Englewood Cliffs, N.J.: Prentice Hall; 1995.
26. Vach W, Roßner R, Schumacher M, (1996). Neural networks and logistic regression: Part II. *Computational Statistics & Data Analysis*. 1996;21(6):683-701.