

## Using Hybrid Data to Model Student Success in Engineering With Artificial Neural Networks

P. K. Imbrie, Joe. J.J. Lin, Ken Reid, and Alexander Malyscheff

*School of Engineering Education, Purdue University, West Lafayette, Indiana, USA*

Phone: (765) 496-7225

Fax: (765) 494-5819

imbrie@purdue.edu, linjj@purdue.edu, kenreid@purdue.edu, and amalysch@purdue.edu

### **Context:**

A number of studies have quantified the problem of high attrition in engineering programs, especially after the first year of study. For example, one large scale study by Astin (1993) involving approximately 25,000 students in over 300 institutions found that less than half of the students who began in engineering graduated with an engineering degree. Another study showed first-year attrition from engineering increased from about 12% in 1975 to 25% in 1990 (Beaufait, 1991). A report issued by the The National Academies (Augustine, 2005) shows undergraduate programs in science and engineering to have some of the lowest retention rates among academic disciplines.

Studies examining students leaving engineering have shown that those who persist and those who leave engineering are quite similar in cognitive ability and scholastic achievement, (Astin et. al., 1992) but tend to show differences in noncognitive characteristics (Jagacinski, 1981). Models designed to predict retention using multiple measures of cognitive and noncognitive attributes show the greatest promise of predicting student retention (e.g., Besterfield-Sacre et. al., 1997).

Therefore, the goal of this study was to develop a multi-measure artificial neural network (NN) model capable of predicting student success after the first year. Such a model has the potential to help faculty and advisors provide informed course selection advice to beginning engineering students and/or could facilitate the development of intervention programs to improve retention. The NN model used a combination of a student's cognitive and noncognitive ability as input and student success after the first-year as output. For the purpose of the paper "student success" was operationalized as retained in engineering. For this investigation student data was collected at a large Midwestern University prior to the start of the students first year during 2004 ( $N=1595$ ), 2005 ( $N=1814$ ) and 2006 ( $N=1838$ ). These data were used as input to artificial neural network (NN) models developed to predict students' retention.

### **Research Questions:**

- 1) Is the instrument used to assess student noncognitive characteristics psychometrically reliable and stable over time?
- 2) What is the predictive accuracy of a neural network model designed to predict student retention given students' non-cognitive and cognitive data as inputs?

### **Theoretical Framework:**

Numerous studies from both engineering and elsewhere have identified various factors that potentially influence student success in college. For this investigation the authors specifically selected factors consistent with the literature that institutions can affect during a student's first year. As such, nine noncognitive constructs were identified for use as inputs to the model. They include:

*Motivation:* Motivation was evaluated using the Academic Intrinsic Motivation Scale (AIMS) (French and Oakes, 2003), consisting of 25 items with four subfactors.

*Metacognition:* The Metacognition scale consists of four subfactors, describing a student's perception of their strategies for monitoring and modifying their cognition (Pintrich, 1999) (O'Neil and Abedi, 1996).

*Deep learning and Surface learning:* Items for these scales were adapted from the Study Process Questionnaire (SPQ) (Biggs et al., 2001). Each consists of two subfactors.

*Self-efficacy:* Many studies indicate the importance of self-efficacy and it has been shown to be predictive of retention (Bandura, 1993) (Pajares, 1996). No subfactors were found in this scale developed specifically for this instrument.

*Expectancy-value:* This scale was developed based on work from Wigfield and Eccles (2000) and includes four subfactors.

*Major indecision:* The scale items in major indecision were developed based on models of career indecision as described by Osipow (1999). The Major Indecision scale consists of four subfactors.

*Leadership:* The student's self assessment of their leadership is based on four subfactors. Characteristics of leadership are theorized to have a positive effect on student retention (Hayden, Holloway, 1985).

*Team vs. individual orientation:* This construct was developed specifically for this instrument and is comprised of ten items and two subfactors.

The instrument assessing these nine noncognitive constructs consists of 161 5-level Likert scale items. Eleven cognitive items also collected in this study include overall grade point average (GPA) and core GPA from high school, standardized test results (SAT/ACT), average high school grades in mathematics, science, English classes and the number of semesters taking these classes.

## **Methodology:**

Data analysis is based on incoming freshman engineering students from a large Midwestern university during the 2004 ( $N=1595$ ), 2005 ( $N=1814$ ) and 2006 ( $N=1838$ ) academic years. In the aggregate population, 17.0% were female, and 83.0% were male. Ethnicity was as follows: 2.17% African American, 0.48% American Native, 9.44% Asian/Pacific Islander, 2.88% Hispanic, 78.21% Caucasian and 6.81% Others.

The survey of noncognitive characteristics and collection of the eleven cognitive characteristics were part of an overall data collection process completed prior to the beginning of the students' first academic year.

Psychometric properties of the scale including examination of means, standard deviations and Cronbach's coefficient alpha were determined. In addition, confirmatory factor analysis (CFA) was used to validate scale constructs and the stability and repeatability of the instrument over time was demonstrated using McDermott's three stage cluster analysis (McDermott, 1998) for each cohort. These constructs were then used as inputs to the predictive neural network models discussed below.

In this study, feed-forward back-propagation neural network models were developed to predict engineering students' retention based on the collected non-cognitive and cognitive factors. Three neural network models with different input datasets were investigated. Model A used a non-cognitive dataset of 60 items selected from factors shown to have the greater effects. These items were extracted from the 161-item survey and represented all constructs except for self-efficacy and team vs. individual orientation. Model B employed the 11 cognitive items (GPAs, standardized test scores, and grades and numbers of semesters in math/science/English). Model C used the combined inputs from Model A and B, with 71 items from both cognitive and non-cognitive datasets as inputs to its model.

In preliminary experiments the numbers of hidden nodes in these three models were determined by sensitivity analysis, using the 2004 dataset and 10-fold cross-validation. The numbers of hidden nodes for the three models were therefore selected as 30, 11 and 36 in models A, B, and C, respectively. Levenberg-Marquardt method was used as the back-propagation algorithm in the training process. Activation functions for the hidden and output layer were both hyperbolic tangent sigmoid functions (Haykin, 1994). The performance function used was mean square error (MSE). These neural network models were implemented using Matlab 7 (Hanselman & Littlefield, 2005).

All models were initially trained and tested using the 2004 cohort data with K-fold cross-validation. The overall prediction accuracy for each of the models was obtained by computing the average of each 10-fold cross-validation result. The true predictive performance of the NN models was evaluated using the 2004 cohort data for training and cross-validation and then tested with students from the 2005 and 2006 cohorts. The accuracy results described below will demonstrate how well the neural network models predict on future students when they were trained with previous student cohort data.

### Findings and Conclusions:

A bifactor structure was demonstrated; this model allows individual items to load to a construct (such as Metacognition) as well as subfactors within that overall domain. Homogeneity of the constructs and subfactors was verified with Cronbach’s coefficient alpha > 0.8 in each case except one (Team vs. Individual Orientation  $\alpha=0.76$ ). In cases of fewer than ten items in a subfactor, the Spearman-Brown formula was used to project alpha to ten items (Alsawalmeh and Feldt, 1999).

The bifactor model structure requires all subfactors to be orthogonal and all items to have nonzero loadings to the main construct (Gibbons and Hedecker, 1992), while taking the conditional dependence of items within a major construct into account (Immekus and Imbrie, 2008). The validity of this structure vs. a unidimensional model was assessed using the chi-square statistic, the Comparative Fit Index (CFI), Goodness of Fit Index (GFI) and Root Mean Square Error of Approximation (RMSEA). In each case, the bifactor model showed an improved and acceptable fit. Table 1 shows that each construct meets levels for acceptable model fit in at least two of the three indices.

**Table 1: Psychometric data and Confirmatory factor analysis fit indices for each construct**

	$\alpha$	$\chi^2$	$\chi^2$ df	pr> $\chi^2$	GFI	RMSEA	CFI
<b>Deep Learning</b>	0.82	84.2	25	<.0001	0.984	0.049	0.972
<b>Surface Learning (revised)</b>	0.79	32.2	25	0.153	0.993	0.017	0.997
<b>Expectancy Value</b>	0.90	1372.5	432	<.0001	0.917	0.047	0.908
<b>Leadership</b>	0.89	443.0	207	<.0001	0.961	0.034	0.961
<b>Major Indecision</b>	0.85	324.2	159	<.0001	0.938	0.046	0.959
<b>Metacognition</b>	0.92	599.3	151	<.0001	0.942	0.055	0.942
<b>Motivation</b>	0.93	999.5	250	<.0001	0.923	0.055	0.927
<b>Team vs. Individual Orientation</b>	0.76	84.1	25	<.0001	0.984	0.049	0.969
<b>Self-Efficacy</b>	0.90	<i>(no subfactors specified for self-efficacy)</i>					
$\alpha$ = Chronbach’s coefficient alpha		GFI = Goodness of Fit Index					
$\chi^2$ = Chi-squared		RMSEA = Root Mean Square Error of Approximation estimate					
$\chi^2$ df = Chi-square degrees of freedom		CFI = Bentley’s Comparative Fit Index					

The stability and repeatability of the instrument over time was demonstrated through cluster analysis. McDermott’s three stage cluster analysis (McDermott, 1998) was used to investigate each

cohort of students, shown to cluster to three distinct clusters. Figure 1 shows the three-cluster solution for each cohort.

In Table 2, Cattell’s similarity coefficients (Cattell, 1978) showed each of the three clusters to be highly correlated (coefficient values > 0.90) with one exception, cluster 3 from 2004 to cluster 3 from 2006, which was shown to be within an acceptable range (coefficient > 0.70). All clusters not expected to be similar were shown to be below 0.70, indicating lack of similarity. These results show the model is stable and comprised of reliable constructs and subfactors.

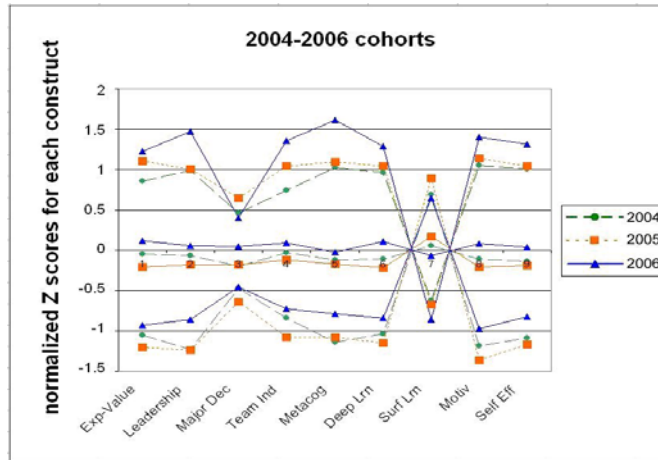


Figure 1: Overlay of results of cluster analysis, 2004-2006 student cohorts

Table 2: Cattell's similarity coefficients, 2004-2006

cluster		2005			2006					2006		
		1	2	3	1	2	3			1	2	3
2004	1	<b>0.99</b>	0.28	0.22	<b>0.97</b>	0.58	0.01	2005	1	<b>0.93</b>	0.66	-0.05
	2	0.47	<b>0.98</b>	-0.36	0.24	<b>0.95</b>	-0.47		2	0.14	<b>0.90</b>	-0.50
	3	0.24	-0.36	<b>0.98</b>	0.47	-0.20	<b>0.84</b>		3	0.37	-0.26	<b>0.90</b>

The prediction performance of the NN models using different cognitive and non-cognitive data as inputs are presented in Table 3. Prediction accuracy measures the percentage of correct predictions within the total number of observations. Probability of detection (POD) for not-retained students measures the percentage of not-retained students who were correctly identified by our system. The results showed that out of the three models, the model combining both cognitive and non-cognitive datasets (Model C) had the best predictive performance. Also, Model A using only non-cognitive inputs appeared to be competitive when compared with the cognitive-only model B. These results confirmed the value of bringing non-cognitive factors into the prediction of student retention, and were consistent with the findings of Besterfield-Sacre et al. (1997).

Table 3: Comparison of prediction results between different NN models (2004 cohort)

Model	Model A	Model B	Model C
Input data of Neural Network model	60 Non-cognitive items	11 Cognitive items	60 Non-cognitive and 11 cognitive items
Prediction Accuracy	<b>70.5%</b>	<b>69.7%</b>	<b>71.3%</b>
POD Not Retained	37.9%	36.2%	40.5%

The authors also performed analyses using NN models trained with 2004 student data to predict retention status of 2005 and 2006 cohorts. In Table 4, the prediction accuracy for 2005 and 2006 cohorts were comparable to the accuracy of predicting 2004 cohorts. This indicates that this neural

network model trained by 2004 data performed consistently when predicting students' status from different cohort years.

**Table 4: Predicting student's retention status in 2005 and 2006 cohorts with NN model trained by student data from 2004 cohort**

Model	Model A		
Cohort data used in training	2004		
Cohort data used in testing	2004	2005 (N=1814)	2006 (N=1838)
Prediction Accuracy	<b>70.5%</b>	<b>70.3%</b>	<b>71.8%</b>
POD Not Retained	37.9%	37.4%	40.2%

**Recommendations:**

With attrition around 20% of freshman students after one year and 50% retention after four years, early preventive intervention from academic counselors is an important way to help students who may consider leaving engineering. However, it is difficult to identify students with a high risk of leaving engineering early enough to make any intervention meaningful and effective. This instrument and the predictive neural network model will contribute to the development of individual intervention strategies and improve the chances for students to persist and succeed in engineering.

The instrument described here is psychometrically stable and reliable. However, additional noncognitive constructs may increase the accuracy of the predictive model. Additional constructs have been included in the most recent version of the instrument: these include self-worth, expectations of social climate, implicit beliefs, goal orientation and intent to persist. Future efforts on the NN models will focus on improving the prediction accuracy by using methods such as fuzzy logic techniques or support vector machines.

**Acknowledgements:**

The researchers wish to acknowledge the support provided by National Science Foundation, Division of Engineering Education and Centers (Award No. 0416113).

**References:**

Alsawalmeh, A. and L. Feldt, (1999). Testing the Equality of Two Independent  $\alpha$  Coefficients Adjusted by the Spearman-Brown Formula, *Applied Psychological Measurement*, Vol. 23, No. 4, pp. 363-370.

Astin, A., T. Banta, P. Cross, E. El-Khawas, P. Ewell, P. Hutchings, et al., (1992). Principles of Good Practice for Assessing Student Learning, *American Association for Higher Learning*, 1992.

Astin, A. W. (1993). Engineering Outcomes, *ASEE Prism*, September 1993, pp. 27-30.

Augustine, N. (chair) (2005). National Academies Committee on Prospering in the Global Economy of the 21st Century, *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*, National Academies Press, Washington, D.C., 2005.

Beaufait, F. W. (1991). Engineering Education Needs Surgery, *Frontiers in Education Conference*, 1991

Besterfield-Sacre, M., Atman, C. and Shuman, L. (1997). Characteristics of Freshman Engineering Students: Models for Determining Student Attrition in Engineering, *Journal of Engineering Education*, vol. 86, pp. 139-149.

Biggs, J., D. Kember and D. Leung (2001). The Revised Two-Factor Study Process Questionnaire: R-SPQ-2F, *British Journal of Educational Psychology*, Vol. 71, pp. 133–149, 2001.

Cattell, R. B. (1978). *The Scientific Use of Factor Analysis in Behavioral and Life Sciences*, New York, NY, Plenum Press, 1978.

- French, B. and W. Oakes (2003). Measuring Academic Instinsic Motivation in the First Year of College: Reliability and Validity Evidence for a New Instrument, *Journal of the First-Year Experience*, Vol. 15, No. 1, pp. 83-102, 2003.
- Gibbons, R. and D. Hedeker (1992). Full-Information Bifactor Analysis, *Psychometrika*, Vol. 57, pp. 423-426, 1992.
- Hanselman, D. C., & Littlefield, B. (2005). Mastering MATLAB 7. Upper Saddle River, NJ: Pearson/Prentice Hall.
- Hayden, D. and E. Holloway (1985). A Longitudinal Study of Attrition among Engineering Students, *Engineering Education*, Vol. 75, No. 7, pp. 664-668, 1985.
- Haykin, S. S. (1994). Neural networks: a comprehensive foundation. New York: Macmillan College Publishing Company
- Jagacinski, C. and LeBold, W. (1981). A Comparison of Men and Women Undergraduate and Professional Engineers, *Journal of Engineering Education*, Vol. 72, pp 213-220, 1981.
- McDermott, P.A. (1998). Megacluster Analytic Strategy for Multistage Hierarchical Grouping with Relocations and Replications, *Educational & Psychological Measurement*, Vol. 58, pp. 677-686, 1998.
- O'Neil, H. and J. Abedi (1996). Reliability and validity of a state metacognitive inventory: Potential for alternative assessment, CSE Technical Report 469, National Center for Research and Evaluation, Standards, and Student Testing (CRESST), University of California, Los Angeles, December 1996.
- Osipow, S. H. (1999). Assessing Career Indecision, *Journal of Vocational Behavior*, Vol. 55, No. 1, pp. 147-154.
- Pintrich, P. R., and E. V. DeGroot (1999). Motivational and self-regulated learning components of classroom academic performance, *Journal of Educational Psychology*, Vol. 82, pp. 33-40, 1999.
- Wigfield, A. and J. Eccles (2000), Expectancy–Value Theory of Achievement Motivation, *Contemporary Educational Psychology*, Vol. 25, pp. 68–81, 2000.