International Journal of Engineering Education Vol. 30, No. 2, pp. 360–377, 2014
Printed in Great Britain

0949-149X/91 $3.00+0.00

Defining the Majors that Comprise "STEM": An Analytical Method for Looking Beyond the Classical Acronym*

GILLIAN M. NICHOLLS
The University of Alabama in Huntsville, Industrial & Systems Engineering & Engineering Management, Huntsville, AL, 35899, USA.
E-mail: gillian.nicholls@uah.edu

HARVEY WOLFE, MARY BESTERFIELD-SACRE and LARRY J. SHUMAN
University of Pittsburgh, Department of Industrial Engineering, Benedum Hall, 3700 O'Hara Street, Pittsburgh, PA 15261, USA.
E-mails: Hwolfe@pitt.edu, mbsacre@pitt.edu, shuman@pitt.edu

There is no universally agreed-upon definition for "STEM." Research in STEM education would be aided by developing a convention for how to categorize students' potential college majors. This study develops and tests a definition of STEM and a method for verifying the definition's validity and usefulness. A classification scheme is developed from reviewing prior educational research; examining the National Science Foundation's categorizing of fields of study; and constructing a predictive model. Demographic, attitudinal, experiential, and standardized test score variables from the National Educational Longitudinal Study of 1988 (NELS:88) dataset are used to predict which students will earn a STEM degree vs. another educational outcome. The predictive model tests the classification of majors into different four year degree outcomes. It tests the inclusion of majors beyond the classic "bench" sciences, engineering, and mathematics. Models were fitted to predict whether students had a STEM vs. other outcome, and the predicted outcomes were compared to actual outcomes. The models with more divergent outcomes had greater accuracy. Students earning a degree involving significant quantitative coursework in a field not generally considered STEM had more factors in common with students earning a Non-STEM degree. A narrow definition of STEM focusing on the bench sciences, engineering, and mathematics produces more accurate predictions. This method for classifying educational outcomes allows researchers to precisely define which majors or non-degree outcomes are included in the analysis. The method is flexible enough to accommodate a mutually exclusive and exhaustive set of outcomes.

Keywords: STEM definition, STEM persistence; NELS:88 Longitudinal Study; Logistic Regression modeling

1. Introduction

The rising importance of education in engineering and the sciences has led to the growth of engineering education as a discipline with its own research initiatives. Grasso, Callahan, and Doucett [1] discussed the need for an interdisciplinary approach to teaching engineering that integrates it not only with the bench sciences but also the social sciences and the humanities. Splitt [2] reached a similar conclusion arguing that engineering education was in need of reform that would produce graduates capable of designing engineering systems that not only solve technical problems but also address constraints in the wider society including environmental, sustainability, ethical, and economic concerns. Benson, Becker, Cooper, Griffin, and Smith [3] reviewed the history and development of the engineering education discipline and observed that research in this field can provide the evidence-based findings to fundamentally improve how engineers and scientists are taught.

Engineering educational research is often pursued under the larger umbrella of STEM education. Research in STEM vs. Non-STEM education is complicated by a lack of consistency in how these terms are defined. Without a standardized set of definitions for STEM and Non-STEM it is difficult to directly compare findings from different research studies or even determine how to utilize their results. Extending the body of knowledge in STEM vs. Non-STEM education requires a standard definition of what college majors and degrees earned constitute STEM. This is a necessary research topic because the educational literature has warned for some time that the United States is at risk of losing its place as a major producer of engineering graduates and new technologies. There is a need for the field of engineering education to adapt to the global marketplace [4]. The volume of Chinese and Indian engineering graduates has increased sharply in recent years, but there is not a consistent way to compare engineering degrees among the various nations [5, 6]. The international competitiveness of the U.S. educational system is an ongoing discussion. To maintain its relative position in the world, the U.S. needs to educate the engineers, scientists, and mathematicians who will develop the next generation of innovations. The President's Council of Advisors on Science and Technology (PCAST) [7] has concluded that economic forecasts indicate approximately one million

* Accepted 5 November 2013.

additional STEM degree holders must be graduated in the next decade than were previously estimated. Improving retention of STEM students from 40% to 50% would address the majority of this forecasted requirement. Increasing the number of students within the engineering "pipeline" is another approach to address the shortage in engineering graduates. Students engaged in the broader field of STEM are good candidates for recruitment into engineering.

In reviewing the literature, it is apparent that there is no consistent scheme for classifying college majors into STEM vs. other Non-STEM categories. Prior research of persistence in various degree programs often does not make it clear how STEM is defined. Many of the papers examining STEM education refer only to majors in "Science," "Engineering," "Mathematics," or some combination of these fields. In some studies, "Science" is limited to just the bench sciences such as chemistry and biology while in other studies majors such as psychology and sociology are also included. Medical and health care majors are classified as STEM by some researchers and not by others. In a few cases a comprehensive list is provided to indicate which majors and degrees are studied as STEM outcomes. The literature in general is not sufficiently detailed to understand what is included in STEM. Where sufficient detail is provided there is little agreement. For analytical models to be useful, a standard definition of STEM is required.

This examination of the definition of STEM was undertaken as a pre-requisite for developing a model that would explore factors about students and predict whether they would ultimately achieve a college degree in a STEM major or not. Projecting students' progression towards a final educational result requires a means of differentiating between outcomes. This is necessary to analyze records and construct models to identify those significant factors that differentiate students who earn a STEM degree from those who do not.

Crafting a standard definition of STEM began by considering the aspects of college majors that would or would not qualify them to be considered STEM. Then the majors were placed in an initial categorization based on the formal name of the major, the body of coursework that is presumed to accompany that major, and the anticipated career application of a degree in that field. These classifications were compared to those suggested by prior researchers, and adjusted if a persuasive case could be made for a particular academic field to be considered STEM or other. Then statistical tests were conducted to determine if there were significant differences among the classifications. The results indicated that a sharply focused definition of STEM produced predictive models with greater accuracy. The characteristics of students who majored in the classical STEM areas of the bench sciences, engineering, and mathematics were significantly different than those for students in other majors including some which required extensive scientific and mathematics coursework.

The potential applications of this research include analysis of undergraduate students' entrance into, progression through, and graduation from STEM majors. Among the important research issues that a precise definition of STEM would aid are measuring the effectiveness of recruiting strategies; the impact on STEM students of different academic policies; the time points of greatest risk for not persisting; the potential for retaining students within STEM if they choose to switch to a different STEM major; and the time to graduation within the different STEM fields. In particular, a rigorous definition of STEM that allows researchers to predictively model capable STEM students at risk of switching out due to academic difficulties or waning interest can enable delivery of pro-STEM interventions to increase retention.

2. Prior research classifying college majors

Some prior researchers have assembled a set of majors that were classified as STEM or a subset of STEM based on their individual research interests and their interpretation of the relevant academic literature. For example, Adelman [8] created a list of majors that were classified as Science or Engineering for his 1998 analysis of the paths taken by students in their undergraduate engineering careers. One of his key points in separating engineering from science was that the practice of engineering involved working closely with clients to satisfy their expectations. This involved far more social interaction and required greater awareness of customer service issues than the practice of a bench science. He indicated that it was important to understand that science and engineering students can be quite different.

Seymour and Hewitt [9] defined a set of majors as science/mathematics/engineering (SME) in their research. For the purposes of this analysis SME and STEM are synonymous. The broadly defined majors considered to be STEM were mathematics/statistics, physical sciences, biological sciences, engineering, and agriculture. The broadly defined Non-STEM majors were computer science, health, business, education, all humanities and fine arts, other non-technical, and undecided. The fields of study included in these majors by Seymour and Hewitt are listed later in the paper. It is interesting to note that computer science was classified as Non-

STEM even though the majority of computer science departments are within colleges of engineering rather than arts and sciences. This decision by Seymour and Hewitt highlights the classification problem for researchers.

Smyth [10] classified certain majors as SME for his analysis of ethnic differences in graduation from selective colleges with a science degree. This built upon earlier logistic regression analysis of graduation trends by race/ethnicity at selective colleges by Smyth and McArdle [11]. Smyth analyzed data obtained from a set of colleges within the College and Beyond (C&B) database [12], developed by the Andrew W. Mellon Foundation (AMF) [13] with data from 34 colleges. Smyth worked with a subset of the C&B colleges chosen by Bowen and Bok [14] as possessing greater academic prestige and that Smyth defined in his analysis as selective in their acceptance of students. Smyth obtained data from the Cooperative Institutional Research Program (CIRP) [15] for 24 of the universities within the C&B database. The other C&B database colleges did not have corresponding CIRP data available. Two of the CIRP variables concerning the students' intended majors and how they rated the importance of contributing to the body of scientific knowledge were utilized in examining differences in SME graduation rates between Caucasian and African-American students. The majors that Smyth categorized as SME (STEM) are compared to those of Seymour and Hewitt in Table 1.

Smyth referred to research by Astin and Astin [16] and Hilton, Hsia, Solorzano, and Benton [17] in his decision to exclude the social sciences and psychology from the STEM category. Prior research had found that very few of the students that declared an intention to major in these two areas switched to a major generally considered as STEM despite approximately half leaving their original majors. This suggested that the students who selected the social sciences and psychology as a major tended to have less academic interests in the standard STEM majors than other potential majors. The accuracy of the models that included social sciences majors to predict STEM graduation was considerably worse than that of regression analyses excluding these majors. When the social sciences and psychology were excluded from the STEM category, the predictive ability of the regression models was more accurate.

In 2008, the State Educational Technology Directors Association (SEDTA) [18] proposed the definition "STEM refers to the areas of science, technology, engineering, and mathematics. STEM initiatives started as a way to promote education in these related areas so that students would be prepared to study STEM fields in college and pursue STEM-related careers. Schools with a strong emphasis on STEM education often integrate science, technology, engineering, and mathematics into the entire curriculum." Tsupros, Kohler, and Hallinen [19] developed a definition of STEM as "STEM Education is an interdisciplinary approach to learning where rigorous academic concepts are coupled with real world lessons. Students apply science, technology, engineering, and mathematics in contexts that make connections between school, community, work, and global enterprise." Lantz [20] did not offer a formal definition but made the case that STEM is trans-disciplinary across all four areas of the acronym rather than simply interdisciplinary. Brown, Brown, Reardon, and Merrill [21] conducted a study that surveyed teachers and administrators about their perceptions of STEM in order to assess the current understanding of what STEM education is. The results of surveying 172 individuals indicated that there isn't a clear agreement of what STEM education is even among professionals engaged in it. Merrill [22] defined STEM education as "a standards-based, meta-discipline residing at the school level where all teachers, especially science, technology, engineering, and mathematics (STEM) teachers, STEM teaching and learning focuses on authentic content and problems, using hands-on, technological tools, equipment, and procedures in innovative ways to help solve human wants and needs." Kelley [23] examined the history of technology education and its connection to STEM education. He concluded that technology is an essential part of the STEM movement and a strong research agenda should be pursued vigorously to advance the overall goals of STEM education.

Barakos, Lujan, and Strang [24] reviewed the literature to discuss various different definitions for STEM and concluded that there isn't a single best answer. Instead they suggested that institutions engaged in STEM education should carefully consider whether there is an agreed-upon interpretation of STEM education for their stakeholders and then develop the educational strategies to support that definition. Gerlach [25] reached a similar conclusion stating that a single definition of STEM wasn't possible because "it means so much for so many different groups of people." Instead he suggested focusing on whatever problems existed in a particular industry, learning about them, solving them, and continuously pursuing innovation. Lewis [26] concluded that efforts to spread engineering education to K-12 schools have been hampered by the lack of an encompassing knowledge base and teachers familiar with the subject matter. Working partnerships between engineers and K-12 instructors are needed to overcome this. The National

Research Council (NRC) [27] concluded that the social sciences would be problematic to include in a definition of science education for K-12 education. Koonce, Zhou, Anderson, Hening, and Conley [28] attempted to address the lack of a consistent definition for STEM by examining the different definitions of STEM in use by STEM researchers, institutions, universities, and organizations. They collected a wide range of data from different STEM entities and created a probabilistic coding scheme that assessed how frequently a particular program of study was included in STEM. They concluded that STEM is defined differently by organizations that focus on education vs. workforce requirements. Inconsistencies in the level of detail about the disciplines included in higher level classifications limited the analysis.

Kokkelenberg and Sinha [29] examined student success in STEM degree programs after excluding the social sciences from the STEM classification. Students that developed an interest in STEM in early high school; who declared their major as a freshman; that possessed strong mathematics skills; and were Asian in race/ethnicity were far more likely to remain in STEM to earn a degree. Nicholls, Wolfe, Besterfield-Sacre, & Shuman [30] examined student persistence to degree attainment in STEM fields using a similar classification scheme for STEM. Students' probability of earning a STEM degree was predicted based on eighth grade demographic, attitudinal, and academic performance. Proficiency in mathematics and science, high SAT/ACT math scores, high parental expectations, high personal academic expectations, and a racial category of either Asian or African-American were found to increase the probability of earning a STEM degree.

The National Science Foundation (NSF) has created an official list of which majors it considers science, engineering, and health related. NSF issues annual figures of degrees earned at the bachelors, masters, and doctoral levels by major, gender, and citizenship within the United States [31] through statistical reports available on its website. NSF classifies majors as falling within the science, engineering, or health fields. An extensive taxonomy of fields of study has been developed by NSF to aid in the classification process [31]. The broad categories included under the science heading are agricultural sciences, biological sciences, computer sciences, Earth/atmospheric/ocean sciences, mathematical sciences, physical sciences, psychology, and social sciences. The social sciences include fields such as economics, political science, sociology, linguistics, anthropology, archeology, criminology, and geography. The engineering category includes aerospace engineering, chemical engineering, civil engineering, electrical engineering, industrial engineering, mechanical engineering, metallurgical/materials engineering, and a general other set for smaller engineering disciplines. The health fields include medicine, dentistry, veterinary medicine, health systems/service administration, nursing, pharmacy, and rehabilitation/therapeutic services.

The National Center for Education Statistics (NCES) [32] has not created an "official" definition of which majors constitute STEM. NCES collects data about the educational progress of students and makes several longitudinal datasets available to researchers. Its function is to collect and disseminate statistics about education rather than to specifically classify portions of education for research purposes. For example, NCES has conducted a series of longitudinal studies of students' progression through high school and college. The most recent study that has been completed is the National Educational Longitudinal Study of 1988 (NELS:88) [33] which began with students in eighth grade in 1988 and continued over a twelve year period with four follow up waves of data collection in 1990, 1992, 1994, and 2000. NCES is currently conducting the Education Longitudinal Study of 2002/06 (ELS:02/06) [34] which began with students in tenth grade and follows them through high school and into college or the labor market. Another round of follow up interviews was conducted in 2012 to determine persistence and attainment in postsecondary education.

However, NCES through Chen and Weko [35, p. 2] did examine the question of which fields could be considered STEM for the purpose of studying entrance into and persistence through a STEM major. Students entering STEM tended to be younger, traditional students; had parents with a higher education level; and had demonstrated strong mathematics skills. More males than females pursued STEM; almost half of Asian/Pacific Islanders chose STEM majors; and a higher proportion of foreign students vs. domestic students entered STEM. Chen and Weko briefly discussed the broad definition employed by NSF as well as the frequent exclusion of the social/behavioral sciences from some STEM research. They chose to define STEM as including the natural sciences, engineering, engineering technologies, mathematics, and computer/information sciences while excluding the social/behavioral sciences. The physical and biological/agricultural sciences were included within the natural sciences category. NCES [36] also examined STEM students in postsecondary education using the same classification of majors that NSF had although the social/behavioral sciences, humanities, business, education, and health sciences were examined separately as "selected Non-STEM fields."

3. Methodology

3.1 Expansive vs. narrow definition of STEM

This study examines whether the definition of STEM should be formally expanded to include majors beyond those implied by the STEM acronym: science, technology, engineering, and mathematics, and if so, what should be included. There were two schools of thought explored in prior research. The first approach was to define STEM narrowly along the lines used by Smyth and Seymour and Hewitt. This approach limits STEM to the ''hard'' sciences, the engineering majors, and mathematics while specifically excluding the health fields, virtually all technology majors, and the ''soft'' sciences such as psychology and social sciences. Under this definition, the ''hard'' sciences, the engineering majors, and mathematics are classified as STEM majors and every other college major is classified as ''Non-STEM.'' A second approach was to include the other fields of study that NSF classifies as part of the sciences such as psychology and the social sciences.

Here we consider adding disciplines generally not included in prior research. The question of whether a narrow or expansive definition of STEM is warranted is explored by proposing a clearly delineated classification scheme and conducting statistical analyses to determine if there are significant differences among students in different classes. The logic of developing a more expansive definition of STEM reflects that an increasingly larger number of majors involve significant quantitative coursework in order to prepare for demanding careers that require applying scientific knowledge, mathematical skills, and independent judgment. Under an expanded definition of STEM many of the health fields such as medicine, dentistry, advanced nursing, etc. are included. Psychology and social sciences are included as well since sophisticated statistical analyses are often used to evaluate data in these fields. In addition, many of the business majors are also eligible since business administration, business finance, operations management, and marketing may each involve significant analytical coursework to prepare students for applying quantitative techniques in making business decisions. Consequently students in these majors may take extensive coursework in science and mathematics/statistics so they can apply their knowledge in later careers. For example, an accounting major requires mathematics knowledge and ability. Sophisticated financial models for predicting the results of investments are developed by people with degrees in business finance. Researchers in the fields of psychology and sociology use complex statistical models to evaluate hypotheses. Medical doctors, dentists, and veterinarians take extensive coursework in topics such as biology, chemistry, and math in order to apply this knowledge in treating their patients. Nurses also require a strong background in biological science and chemistry. Medical professionals need the skills to integrate technical data and quantitative knowledge with independent judgment. If these majors require extensive coursework in the areas of majors traditionally classified as STEM, there is a case to be made for considering them STEM majors.

A counterargument in favor of not accepting an expanded definition of STEM is that analytical studies attempting to predict between STEM and Non-STEM are enhanced by having sharp divisions between the outcomes. By ''fencing off'' a narrow definition of STEM that is easy to enforce in the classification of student records, the accuracy of resulting predictive models may be improved. Including records from students with a degree in other fields may blur those divisions and weaken the predictive value of the model as was reported by Smyth [10].

Part of the challenge in determining which of the other potential STEM majors should be included as STEM concerned evaluating the degree obtained. A bachelor's degree in business finance at a quantitatively oriented university may involve significantly more demanding quantitative coursework than at other universities with a more qualitative orientation.

Another argument against the more expanded definition of STEM lies in the way the bachelor's degree is applied. Engineers, scientists, and mathematicians learn quantitative material in order to apply it in a creative way. It is not enough for them to merely apply a difficult quantitative technique; they must understand the technique comprehensively so that it can be adapted as the situation warrants. In contrast, an accountant uses math more as an off-the-shelf tool. A medical doctor treating patients is applying his or her knowledge of biology, chemistry, mathematical facts, and the individual patient's medical history to prescribe an accepted standard treatment. A physician's assistant would need to understand medical terminology and be able to evaluate test results to record medical findings, but the physician's assistant would not be applying technical skills in a fully independent fashion. It is indisputable that psychologists, sociologists, medical doctors, and financial analysts may do groundbreaking research, but this is not the focus of the vast majority of professionals in these fields. Therefore, including degrees earned in these fields as STEM outcomes may cloud the analysis of differences between STEM and Non-STEM students. The blurring of differences would correspondingly degrade the ability of analytical models to accurately predict STEM vs. another outcome.

To reconcile the issues of achieving good predictive accuracy with recognizing the quantitative rigor of a wider set of majors, a compromise was chosen between the narrow and expansive definitions of STEM. Majors in the "hard" or "bench" sciences, engineering, and mathematics were categorized as STEM. Majors other than the "hard" or "bench" sciences, engineering, and mathematics that require extensive quantitative coursework were placed in a third category referred to as "STEM-Related." The remaining majors were categorized as Non-STEM.

The advantage of this classification is that it offered a way to reflect the advantages of both the narrow and expansive definitions of STEM. The value of retaining STEM-Related as a separate category was objectively tested by statistical analysis of potential significant differences among students in each of the three classifications. If a predictive model discriminated between the potential outcomes with acceptable accuracy then there was merit in keeping three categories. If there was insufficient accuracy in discrimination between two of the categories this suggests the compromise STEM-Related category is not useful. Prior research by Nicholls, Wolfe, Besterfield-Sacre, and Shuman [37] has clearly indicated that significant statistical differences exist between STEM and Non-STEM majors, so it was informative to test whether the proposed STEM-Related category is significantly different from STEM and/or Non-STEM.

If analysis found the STEM-Related category to be significantly different from one of the main two categories and not significantly different from the other this would suggest the majors within it could be grouped with the latter category. Then additional statistical tests would be appropriate to identify any significant differences between STEM + STEM-Related vs. Non-STEM or STEM vs. Non-STEM + STEM-Related. If analysis found the STEM-Related category to be significantly different from both of the main two categories then it would remain a separate grouping.

The compromise classification scheme is more complicated, but it offers the prospect of creating a definition of STEM that can be logically tested and evaluated for future applications. Therefore, this method of defining STEM was the one selected for further analysis. We propose the following definition of STEM:

> STEM is a path of study that involves significant coursework in advanced Science, Technology, Engineering, or Mathematics such that successful students acquire a comprehensive understanding of these subjects in order to extend and create knowledge. A STEM career requires extensive quantitative skills that can be utilized creatively and with a high degree of independent authority.

3.2 Categorizing college majors

3.2.1 Selecting majors for the three categories

Student outcomes that resulted in earning a bachelor's degree were divided into three mutually exclusive classes: STEM, Non-STEM, and STEM-Related. Since all the prior research into technical education subjects was consistent in classifying the "hard" or "bench" sciences, engineering, and mathematics as "STEM," this narrow set of majors was automatically placed within the STEM category for this analysis. Computer and information sciences were also categorized as STEM.

The Non-STEM category was applied to majors that clearly did not require extensive coursework in quantitative, technical subjects. This category included the fine arts, English, and other humanities.

The STEM-Related category contains those majors that involve extensive quantitative coursework and represent a potential "gray" area between STEM and Non-STEM. This included the health professions (medicine, dentistry, veterinary, pharmacy, nursing, and clinical therapies), agriculture, forestry, social sciences, psychology, business (accounting, business administration, finance, marketing, and management), and technical fields such as computer programming.

The NELS:88/00 dataset was chosen for this study because it offered a very comprehensive set of variables about students' educational process as well as a large number of records. A total of 12,140 students participated in the final wave of data collection. The longitudinal time span covered (1988 to 2000) meant that it was possible to discern the final college degree outcome of the vast majority of students that pursued such a degree. Table 1 lists the college majors within the NELS dataset classified for this study as STEM, STEM-Related, and Non-STEM and compares these with the conclusions of Seymour & Hewitt, Smyth, Chen & Weko, and the National Science Foundation (NSF) classification of majors. It should be noted that the categorization as STEM or Non-STEM shown for Seymour & Hewitt, Smyth, and Chen & Weko are based on our interpretations of their STEM major classifications in prior published works. The categorizations shown for NSF are based on our interpretations of its annual classification of programs into Science, Engineering, or other fields of study. It is understood that individuals and universities have tailored degrees and joint degrees such that it may appear that multiple classifications apply. However, the list in Table 1 reflects an attempt to provide a precise, reliable means of classifying outcomes based on the most quantitative degree designation. For instance, consider a student that earned a dual degree with one in a STEM field and one in a Non-STEM field. According to these rules of classification, the STEM degree would be primary. It should also be noted that academia has expanded to include new fields of inquiry that were not represented in the NELS dataset. Table 1 is based on college majors that are included in the NELS dataset.

The classification used by the National Science Foundation is very robust in order to accommodate various funding interests. NSF considers all programs that are part of the social sciences field as science. The social sciences include American studies, folklore, linguistics, women's studies, political science, archaeology, geography, and criminology. While practitioners in these fields may use sophisticated quantitative methods in conducting research, the primary focus of most degree-holders would be more qualitative. The inclusion of these fields in NSF's STEM classification is a different approach than other researchers have taken. It is a very inclusive definition of STEM which encompasses the traditional STEM fields of science, technology, engineering, and mathematics as well as the broader social sciences and psychology that are funded by NSF. As such it doesn't directly lend itself to the type of research issues that researchers studying STEM in particular, would examine. Much of the focus in STEM research is directed towards understanding how the pipeline of STEM students can be increased and these fields are more peripheral to STEM. Thus the NSF classification isn't directly applicable for most STEM researchers. However, while the aggregate NSF degree data includes these other fields, the detailed data by field is useful. NSF provides data that allow a researcher to determine the number of bachelors, masters, or doctoral degrees earned in each field so that particular subfields may be included or excluded as desired.

Another aspect of the NSF classification that should be noted is that the medical and health fields are labeled as part of "Health Sciences" but placed outside the category of science and engineering. The decision not to include the health sciences as part of STEM seems inconsistent given the inclusion of many social science fields. Considering the range of topics within the social sciences that were considered to be part of Science and Engineering, the health fields would appear to belong there as well. NSF's reason for this classification may be related to the division by the U.S. Congress of research funding between NSF and the National Institutes of Health (NIH). Some prior research has focused on a more narrow classification of STEM that leaves out much of the social sciences while including some of the health sciences.

The degree categories used by Seymour and Hewitt [9], Smyth [10], and Chen and Weko [35] to classify STEM are quite narrow so that the social science fields are considered Non-STEM. Smyth included several health-related fields of study including medicine, dentistry, and veterinary medicine as STEM while Seymour & Hewitt, Chen & Weko, and NSF did not. NSF classified these fields within the "Health Sciences" area of "Non-Science and Engineering." All of the business finance and management majors were considered to be Non-STEM by NSF, Smyth, Chen and Weko, and Seymour and Hewitt.

One unique aspect of the classification scheme proposed here is the attempt to formally evaluate whether a third category, STEM-Related, is valid for majors that have not historically been considered to be STEM but do involve significant quantitative coursework. This method of classification offers the increase in predictive accuracy found in other studies by having a narrow classification for STEM with the option of expanding the definition of STEM to include additional fields if warranted. The proposed classification extends beyond that of prior research by permitting the examination of a much wider array of possible educational outcomes in a single study.

The main contribution of the methodology presented in this research is the logical procedure for differentiating between educational outcomes which is tested by determining if the majors classified as STEM-Related are more similar to Non-STEM majors than to STEM majors. The methodology can also be applied in scenarios where a researcher wished to explore a different classification scheme. This methodology is applicable to any grouping of college majors. In this analysis, a breakdown that differs from that of NSF is formally tested. A logical next step would be to test a different classification scheme for the educational outcomes.

3.2.2 Creation of additional categories

The classification scheme presented in Table 1 outlines three possible educational outcomes for individuals pursuing a four year college degree: STEM, STEM-Related, and Non-STEM. However, there are other educational outcomes besides earning a four year college degree. The additional categories of "Sub 4-yr Degree" and "No Degree" were added to the classification scheme to ensure a set of mutually exclusive and exhaustive outcomes such that each student outcome would be clearly classified into just one category. The "Sub 4-yr Degree" included those who earned a college degree no higher than an Associate degree, and the "No Degree" included

Table 1. Comparison of STEM vs. Non-STEM Major Classification by Researcher

	Authors’ Classification			Seymour & Hewitt		Fred Smyth		Chen & Weko		NSF	
Major/field of study (as specified by NELS 88/00)	**STEM**	**STEM-Related**	**Non-STEM**	**STEM**	**Non-STEM**	**STEM**	**Non-STEM**	**STEM**	**Non-STEM**	**STEM**	**Non-STEM**
Agriculture		**Y**		Y			Y	Y		Y	
Agricultural science		**Y**		Y			Y	Y		Y	
Natural resources		**Y**			Y		Y	Y		Y	
Forestry		**Y**		Y			Y	Y		Y	
Architecture			**Y**		Y		Y		Y		Y
American civilization			**Y**		Y		Y		Y	Y	
Area studies			**Y**		Y		Y		Y	Y	
African-American studies			**Y**		Y		Y		Y	Y	
Ethnic studies-not Black/area studies			**Y**		Y		Y		Y	Y	
Accounting		**Y**			Y		Y		Y		Y
Business-finance		**Y**			Y		Y		Y		Y
Business-business/management systems		**Y**			Y		Y		Y		Y
Business-management/administration			**Y**		Y		Y		Y		Y
Business-secretarial			**Y**		Y		Y		Y		Y
Business-business support			**Y**		Y		Y		Y		Y
Business-marketing/distribution			**Y**		Y		Y		Y		Y
Journalism			**Y**		Y		Y		Y		Y
Communications			**Y**		Y		Y		Y		Y
Communication technology		**Y**			Y		Y		Y		Y
Computer programming		**Y**			Y	Y			Y	Y	
Data processing technology		**Y**			Y	Y			Y	Y	
Computer and information sciences	**Y**				Y	Y		Y		Y	
Consumer services-cosmetology			**Y**		Y		Y		Y		Y
Consumer services-mortuary			**Y**		Y		Y		Y		Y
Education-early childhood			**Y**		Y		Y		Y		Y
Education-elementary			**Y**		Y		Y		Y		Y
Education-secondary			**Y**		Y		Y		Y		Y
Education-special			**Y**		Y		Y		Y		Y
Education-physical education			**Y**		Y		Y		Y		Y
Education-other			**Y**		Y		Y		Y		Y
Engineering-electrical	**Y**			Y		Y		Y		Y	
Engineering-chemical	**Y**			Y		Y		Y		Y	
Engineering-civil	**Y**			Y		Y		Y		Y	
Engineering-mechanical	**Y**			Y		Y		Y		Y	
Engineering-all other	**Y**			Y		Y		Y		Y	
Engineering technology		**Y**		Y		Y		Y		Y	
Spanish			**Y**		Y		Y		Y		Y
Foreign language-non-European			**Y**		Y		Y		Y		Y
Foreign language-European (not Spanish)			**Y**		Y		Y		Y		Y
Health/allied-dental/medical technology		**Y**			Y		Y		Y		Y
Health/allied-Therapy and mental health			**Y**		Y		Y		Y		Y
Health/physical education/recreation			**Y**		Y		Y		Y		Y
Nursing-nurse assisting			**Y**		Y		Y		Y		Y
Health/allied-general and other			**Y**		Y		Y		Y		Y
Nursing-nursing, post-RN		**Y**			Y		Y		Y		Y
Health-audiology		**Y**			Y		Y		Y		Y
Health-clinical health science		**Y**			Y		Y		Y		Y
Health-dentistry		**Y**			Y	Y			Y		Y
Health-medicine		**Y**			Y	Y			Y		Y
Health-veterinary medicine		**Y**			Y	Y			Y		Y
Nursing-registered nurse		**Y**			Y		Y		Y		Y
Health-health/hospital Administration			**Y**		Y		Y		Y		Y
Health-public health			**Y**		Y		Y		Y		Y
Health-preparatory programs			**Y**		Y		Y		Y		Y
Health-dietetics		**Y**			Y		Y		Y		Y
Textiles			**Y**		Y		Y		Y	Y	
Home economics-all other			**Y**		Y		Y		Y		Y
Health-chiropractic			**Y**		Y		Y		Y		Y
Health-pharmacy		**Y**			Y		Y		Y		Y
Health-optometry			**Y**		Y		Y		Y		Y
Vocational home economics-child care			**Y**		Y		Y		Y		Y
Vocational home economics-other			**Y**		Y		Y		Y		Y
Law-paralegal (includes pre-law)			**Y**		Y		Y		Y		Y
Law			**Y**		Y		Y		Y		Y
Letters-American/English literature			**Y**		Y		Y		Y		Y
Letters-creative/technical writing			**Y**		Y		Y		Y		Y
Letters-all other			**Y**		Y		Y		Y		Y
Liberal studies			**Y**		Y		Y		Y		Y
Library/archival science			**Y**		Y		Y		Y		Y
Biological science-zoology	**Y**			Y		Y		Y		Y	
Biological science-botany	**Y**			Y		Y		Y		Y	
Biological science-biochemistry	**Y**			Y		Y		Y		Y	
Biological science-all other	**Y**			Y		Y		Y		Y	
Mathematics-statistics	**Y**			Y		Y		Y		Y	
Mathematics-not statistics	**Y**			Y		Y		Y		Y	
Military sciences			**Y**		Y		Y		Y		Y
Women’s studies			**Y**		Y		Y		Y	Y	

Table 1. (*Continued*)

Major/field of study (as specified by NELS 88/00)	Authors' Classification			Seymour & Hewitt		Fred Smyth		Chen & Weko		NSF	
	STEM	STEM-Related	Non-STEM	STEM	Non-STEM	STEM	Non-STEM	STEM	Non-STEM	STEM	Non-STEM
Interdisciplinary-environmental studies		**Y**			Y		Y		Y		Y
Interdisciplinary-biopsychology		**Y**			Y		Y		Y		Y
Interdisciplinary-integrated science		**Y**			Y		Y		Y		Y
Interdisciplinary-all other			**Y**		Y		Y		Y		Y
Leisure studies			**Y**		Y		Y		Y		Y
Basic/personal skills			**Y**		Y		Y		Y		Y
Philosophy			**Y**		Y		Y		Y		Y
Religious studies			**Y**		Y		Y		Y		Y
Clinical pastoral care			**Y**		Y		Y		Y		Y
Physical sciences-chemistry	**Y**			Y		Y		Y		Y	
Physical sciences-earth science	**Y**			Y		Y		Y		Y	
Physical sciences-physics	**Y**			Y		Y		Y		Y	
Physical sci-not chemistry/physics/earth	**Y**			Y		Y		Y		Y	
Psychology		**Y**			Y		Y		Y	Y	
Protective services			**Y**		Y		Y		Y		Y
Social work			**Y**		Y		Y		Y		Y
Public administration-not social work			**Y**		Y		Y		Y	Y	
Anthropology/archaeology		**Y**			Y		Y		Y	Y	
Economics			**Y**		Y		Y		Y		Y
Geography			**Y**		Y		Y		Y	Y	
History			**Y**		Y		Y		Y		Y
Sociology		**Y**			Y		Y		Y	Y	
Political science			**Y**		Y		Y		Y	Y	
International relations			**Y**		Y		Y		Y	Y	
City planning			**Y**		Y		Y		Y	Y	
Industrial arts-construction			**Y**		Y		Y		Y		Y
Mechanics-transportation			**Y**		Y		Y		Y		Y
Industrial arts-electronics		**Y**			Y		Y		Y		Y
Mechanics-all other			**Y**		Y		Y		Y		Y
Arts-commercial art			**Y**		Y		Y		Y		Y
Precision production			**Y**		Y		Y		Y		Y
Transportation-air			**Y**		Y		Y		Y		Y
Transportation-not air			**Y**		Y		Y		Y		Y
Arts-design			**Y**		Y		Y		Y		Y
Arts-speech/drama			**Y**		Y		Y		Y		Y
Arts-film arts			**Y**		Y		Y		Y		Y
Arts-music			**Y**		Y		Y		Y		Y
Arts-visual/performing/fine			**Y**		Y		Y		Y		Y
Arts-crafts, folk art, artisanry			**Y**		Y		Y		Y		Y
No major			**Y**		Y		Y		Y		Y
{Don't know}			**Y**		Y		Y		Y		Y
{Refused}			**Y**		Y		Y		Y		Y
{Legitimate skip}			**Y**		Y		Y		Y		Y
{Uncodeable}			**Y**		Y		Y		Y		Y
{Not reached-partial/abbrev interview};			**Y**		Y		Y		Y		Y

those who did not seek or did not complete a college degree. Table 2 summarizes the categories.

In addition, three combinations of the five individual educational outcomes were created to allow for an even wider set of comparisons.

- "Other Degree" was constructed by combining STEM-Related and Non-STEM to reflect all students that earned a four year college degree in a subject other than STEM.
- "Non-Degree" was constructed by combining the Sub 4-Yr Degree and No Degree outcomes to reflect all students that did not earn a four year college degree.
- "All Else" was created by combining the STEM-Related, Non-STEM, Sub 4-Yr Degree, and No Degree outcomes to reflect students that experienced an educational outcome other than earning a STEM degree.

The combination categories allowed predictive models such as Degree vs. Non-Degree, STEM vs. Other Degree, and STEM vs. All Else to be constructed.

Table 2. Classification of Student Outcomes by Category

Category	Characteristics
STEM	Earned a bachelors or a masters degree in a STEM field
STEM-Related	Earned a bachelors or masters degree or in a STEM-Related field, and never earned a degree in a STEM field.
Non-STEM	Earned at least bachelors degree in a field other than STEM or STEM-Related
Sub 4 Year	Earned an associate's degree or a certification.
No Degree	Did not complete a college degree or earn a certification in a field.

3.3 Preparation of the dataset

The version of the NELS dataset utilized in this analysis includes detailed transcript information, standardized test scores, average scores in specific topics, and grade point averages as well as a wealth of demographic, experiential, and attitudinal data. The first step in preparing the dataset was examining the records to determine which of the 12,140 students had participated in all five waves of data collection to ensure that complete data was available for this analysis. There were 11,320 students who responded in all five waves of data collection. The next step was to select a manageable set of variables from the more than 7,000 available so that statistical models could be created to predict between outcomes and test the validity of the classification scheme. A set of 66 variables was selected that had been found to show promise in distinguishing between STEM and Not-STEM students in prior research. Because most of the NELS data were categorical in nature, the 66 variables were re-coded to ensure that the values were strictly ordinal, continuous, or binary in nature.

The third step was applying the outcome classification scheme to the 11,320 student records selected for the study. Each record in the dataset needed to be assigned to one of the mutually exclusive and exhaustive categories. The degrees earned and the majors associated with those degrees were examined. The categories for the majors were compared to ranges of values for the five separate educational outcome categories. Each student's outcome was classified as STEM, STEM-Related, Non-STEM, Sub 4-Yr Degree, or No Degree.

If a student pursued a STEM major and the resultant college degree was at least a bachelor's, the student was categorized as STEM. If at least a bachelor's degree was earned in a major other than STEM, but in a STEM-Related topic, the student was categorized as STEM-Related. If the student earned at least a bachelor's degree in a subject other than STEM or STEM-Related then the outcome was categorized as Non-STEM. If an associate's degree or other less than 4-year degree was earned the outcome was classified as Sub 4-Yr Degree independent of the major. If no college degree was earned, the student's outcome was classified as No Degree. The final outcome of the classification is shown in Table 3. It should be noted that the number of students in each category is rounded to the nearest ten per NCES NELS:88 data security requirements.

Once the records were categorized in this manner it became possible to model these outcomes as dependent variables resulting from a set of demographic, academic, attitudinal, and experiential covariates. The various categories were grouped into pairs so that models were fitted to predict students as having a STEM vs. STEM-Related, STEM vs. All Else, STEM vs. Non-STEM, STEM vs. Other Degree, 4 Yr Degree vs. Non-4 Yr Degree, etc. outcome.

3.4 Testing the definition of STEM

The definition of STEM and the scheme for classifying student outcomes was tested by using logistic regression. Logistic regression is used to model scenarios in which only two outcomes are possible and the analyst wishes to predict the probability of one of the two occurring. Since the outcomes are dichotomous, logistic regression is also referred to as binomial regression reflecting that the underlying probability distribution is binomial as opposed to the normal probability distribution underlying linear regression. The dependent variable for logistic regression is a number in the interval [0, 1] representing the estimated probability that the outcome of interest occurs. The primary outcome of interest in this analysis was earning a STEM degree. Predictions were made by selecting a "cutpoint" value within the interval [0, 1] and classifying records with an estimated probability of a STEM outcome greater than or equal to the cutpoint to have a STEM outcome and those with a lower value to have a different outcome. When logistic regression models are able to find sharp distinctions between the dichotomous outcomes, the models have better predictive accuracy. Thus, when the two outcomes are very different, the model's predictive accuracy is higher. When the distinctions between the outcomes are marginal, the model's predictive accuracy is lower. Constructing a series of

Table 3. Numbers of Students within the Dataset Classified by Group

Category	Number of Students	Included in Other Combination Category			
		All Else	**4 Yr Degree**	**Non-4 Yr Degree**	**Other Degree**
STEM	740	No	Yes	No	No
STEM-Related	1,110	Yes	Yes	No	Yes
Non-STEM	2,040	Yes	Yes	No	Yes
Sub 4-Yr Degree	1,700	Yes	No	Yes	No
No Degree	5,530	Yes	No	Yes	No
Total	11,120	10,380	3,890	7,230	3,150

logistic regression models to predict between the various pairs of educational outcomes permitted the relative strength of predictive accuracy to assess how different the outcomes in each pair were from one another.

Separate models were created to predict the probability of:

- STEM vs. STEM-Related
- STEM vs. Non-STEM
- STEM vs. Sub 4 Year Degree
- STEM vs. No Degree
- STEM-Related vs. Non-STEM
- STEM-Related vs. Sub 4 Year Degree
- STEM-Related vs. No Degree
- Non-STEM vs. Sub 4 Year Degree
- Non-STEM vs. No Degree
- Sub 4 Year Degree vs. No Degree

Additional models were created to predict between STEM and combination outcomes such as "All Else" and "Other Degree" as well as "4 Year Degree vs. Non 4 Year Degree". It was theorized that if the definition of STEM should be expanded to include the STEM-Related degrees there would be little predictive strength in the STEM vs. STEM-Related model and more in the STEM-Related vs. Non-STEM model. Conversely, if the STEM-Related category was more similar to the Non-STEM category, it would argue for a more narrow definition of STEM.

Developing a model to predict between a STEM outcome and a different outcome for a given student involves using data to discriminate between the two potential results and then assessing how the predictions compared to actual results. A valuable tool in assessing the accuracy of the discrimination is Receiver Operating Characteristics (ROC) curve analysis [38-41]. The goal is to increase the accuracy of predictions and decrease the likelihood of false positive prediction or false negative predictions. The prediction accuracy is a tradeoff between sensitivity and specificity. In terms of STEM prediction the sensitivity is the probability of correctly classifying a student as having a STEM outcome. The specificity is the probability of correctly classifying a student as having a "Not-STEM" outcome. Fig. 1 depicts an example of an ROC curve showing the results of predicting between a STEM and All Else outcome.

ROC curves visually depict the tradeoff by plotting sensitivity vs. (1 – specificity) for a range of cutpoint values. The value of (1 – specificity) is the probability of a false prediction of a STEM outcome. So depending on the value at which a student is predicted to have a STEM outcome there is a combination of the probability of a correct prediction and an associated probability of an incorrect prediction. Plotting (1 – specificity) on the horizontal axis and sensitivity on the vertical axis produces a curved graph. Ideally, the ROC curve should resemble a vertical line at a low value of (1 – specificity) which transitions to a horizontal line over the

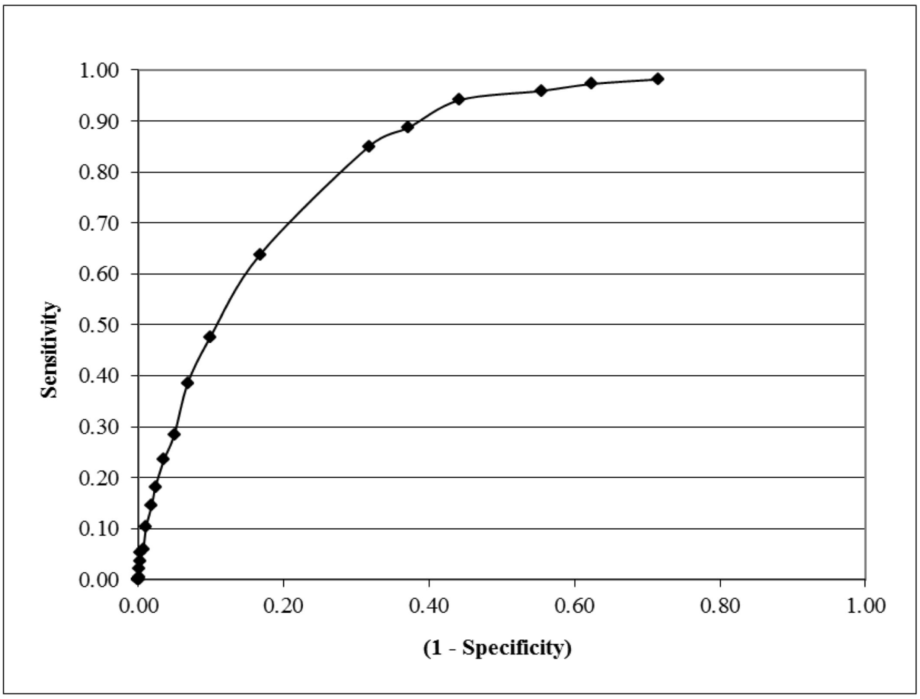

Fig. 1. Example of an ROC Curve depicting the Prediction of a STEM vs. All Else Outcome.

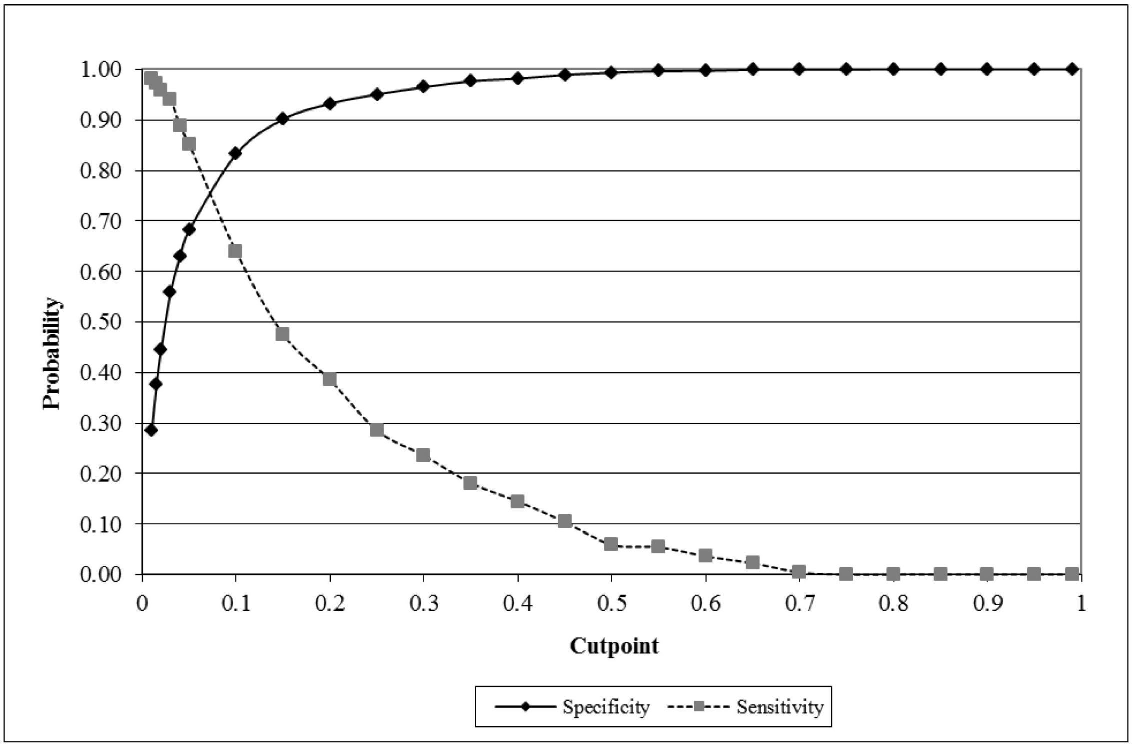

Fig. 2. Example of Sensitivity vs. Specificity by Cutpoint When Predicting Between STEM vs. All Else Outcomes.

remaining values of (1 – specificity). ROC curves with a steep gradient mean that a very high value of sensitivity with a correspondingly low value of (1 – specificity) is attainable. They indicate that a cutpoint can be chosen which offers a high probability of true positive detection and a correspondingly low probability of false positive prediction. If an ROC curve contained the point (0,1) it would indicate the model could be set to a cutpoint that would provide perfect predictive ability. Conversely, a diagonal line from (0,0) to (1,1) would indicate a line of perfect non-discrimination where results were completely random. This diagonal line divides the ROC space with good classification results (better than random) lying above the diagonal and poor results (worse than random) lying below the diagonal. The closer an ROC curve lies to the upper left corner point of (0,1) the better the model's prediction capability, but the distance from the diagonal line in either direction indicates its predictive power.

Plotting the ROC curves for the predictive results of the various models enabled direct visual comparison of the strength of the models and how dissimilar the components of the different two-outcome pairs were. Another measure of the models was the "Area under the Curve" (AUC) value for the logistic regression models. The AUC value is a summary statistic that assesses the predictive strength of the model. In this context, Hosmer and Lemeshow [42, pp. 160–164] indicate that the AUC estimates the probability that a student with a STEM outcome will be predicted to have a higher probability of a STEM outcome than a student with a different outcome.

Figure 2 depicts an example of plotting the sensitivity vs. specificity for the STEM vs. All Else model at a series of different predictive cutpoints. Classifying a student's outcome as STEM vs. another outcome such as Non-STEM is based upon the cutpoint value selected for the prediction threshold. If the cutpoint is set to a low value, then the model will predict more students to have a STEM outcome. As a result more true STEM students will be correctly predicted to have a STEM outcome but correspondingly, more true Not-STEM students will be incorrectly predicted to have a STEM outcome. If the cutpoint is set to a high value, then fewer students will reach that value and be predicted to have a STEM outcome. Thus fewer true STEM students will be correctly predicted to have a STEM outcome and correspondingly fewer true Not-STEM students will be incorrectly predicted to have a STEM outcome. The choice of the cutpoint value determines the combination of sensitivity and specificity in discriminating between the two potential outcomes. The two measures are directly associated for a given cutpoint value. This association means there is a tradeoff between achieving good sensitivity and good specificity in outcome discrimination.

3.5 Validity of the STEM-related category

The logistic regression models developed for STEM vs. STEM-Related, STEM vs. Non-STEM, STEM vs. Other 4 Year Degree, and STEM-Related vs. Non-STEM indicated the hierarchical nature of the models. As theorized, models with more divergent two-outcome pairs had better predictive accuracy than those for which the two outcomes were more similar. Table 4 contains the ROC Curve AUC results of the fitted models for various two-outcome pairs. The models' strength tended to increase as the disparity between the outcomes increased. Hosmer & Lemeshow [42, p. 162] state the predictive accuracy is considered "outstanding" if AUC $\geq$ 0.9, "excellent" if 0.8 $\leq$ AUC < 0.9, "acceptable" if 0.7 $\leq$ AUC < 0.8, and "negligible" if AUC = 0.5.

These results are also found when the outcomes include combinations of categories. For example, the STEM vs. Other Degree model had an AUC of 0.724 which is comparable to the average of the AUC values for STEM vs. STEM-Related and STEM vs. Non-STEM models. The STEM vs. All Else model had an AUC of 0.848 which lies between the STEM vs. STEM-Related/Non-STEM figures and those of STEM vs. Sub 4 Yr. Degree/No Degree. The 4 Year Degree vs. Non 4 Year Degree model predicted student outcomes to be either a 4 Year Degree or a Sub 4 Year Degree/No Degree, and its associated AUC value was 0.882. This reflects a clear divergence between students that did and did not earn a bachelors degree. Thus having an AUC value larger than that of the STEM vs. All Else model is not surprising. The All Else category includes a diverse population of students including those who earned a different 4 year college degree and have more in common with the STEM students than the No Degree students.

One important issue was to determine the validity of creating the STEM-Related category. If this category represented a valid subdivision of the students with bachelor's degrees, significant differences were expected between this category and those of the STEM and Non-STEM categories. Another logistic regression model was fitted to predict between a combination of the STEM and STEM-Related students and their Non-STEM counterparts. This STEM & STEM-Related vs. Non-STEM model had an AUC value of 0.613 which would be considered poor. Evaluation of the STEM & STEM-Related vs. Non-STEM model indicates that it has slightly better predictive accuracy than the STEM-Related vs. Non-STEM model and less accuracy than the STEM vs. Other Degree model. This suggests that while there may be benefit to keeping STEM-Related as a separate category, the students within this category have more in common with the Non-STEM students than the STEM students. Table 5 lists the logistic regression models and the level of predictive accuracy associated with each fitted model.

The comparison of the predictive accuracy for the set of two-outcome models indicates that there are substantial differences between outcomes. In comparing students that had earned a STEM degree with those that had another four year degree, the model's predictive accuracy was rated acceptable because the variables of the two groups were very similar. The variables and their values for the Non-STEM and STEM-Related students had little dis-

Table 4. Hierarchy of Logistic Regression Model AUC Values by Outcome Pair

Outcome	STEM	STEM-Related	Non-STEM	Sub 4 Yr Deg	No Degree
STEM	N/A	0.720	0.743	0.924	0.919
STEM-Related		N/A	0.550	0.885	0.887
Non-STEM			N/A	0.876	0.878
Sub 4 Yr Deg				N/A	0.604
No Degree					N/A

Table 5. Comparison of Logistic Regression Model Accuracy

Model	Predictive Accuracy	Accuracy Scale
STEM vs. STEM-Related	Acceptable	$0.70 \leq AUC < 0.80$
STEM vs. Non-STEM	Acceptable	$0.70 \leq AUC < 0.80$
STEM vs. Sub-4 Year Degree	Outstanding	$AUC \geq 0.90$
STEM vs. No-Degree	Outstanding	$AUC \geq 0.90$
STEM vs. All Else	Excellent	$0.80 \leq AUC < 0.90$
STEM vs. Other 4 Year Degree	Acceptable	$0.70 \leq AUC < 0.80$
STEM & STEM-Related vs. Non-STEM	Poor	$0.50 < AUC < 0.70$
STEM-Related vs. Non-STEM	Negligible	$AUC \approx 0.50$
STEM-Related vs. Sub-4 Year Degree	Excellent	$0.80 \leq AUC < 0.90$
STEM-Related vs. No-Degree	Excellent	$0.80 \leq AUC < 0.90$
4 Year Degree vs. Non-4 Year Degree	Excellent	$0.80 \leq AUC < 0.90$

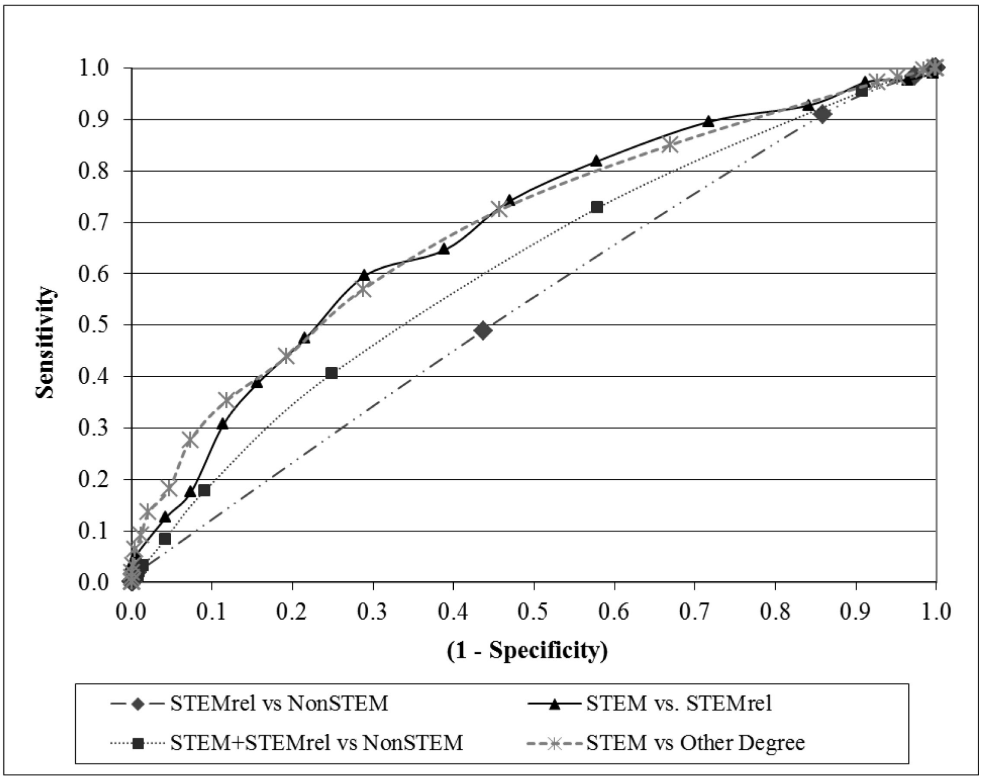

Fig. 3. Comparison of ROC Curves Showing Predictive Strength in Modeling Different College Degree Outcomes.

cernible variation since the model for those two outcomes had negligible ability to differentiate between them. The farther apart the two outcomes were in relation to the ordered set {STEM, STEM-Related, Non-STEM, Sub-4 Year, No Degree} the greater the predictive accuracy of the model. Note from Table 5 that the STEM vs. All Else model had excellent ability to discriminate between the outcomes, suggesting that the mix of other outcomes was sufficiently diverse from the STEM outcome to provide good differentiation between the students. There is probably value in analyzing how re-categorizing the component majors of the other outcomes might affect their differences from the STEM category.

Figure 3 illustrates the ROC curves for the STEM vs. STEM-Related, STEM vs. Other Degree, STEM-Related vs. Non-STEM, and STEM & STEM-Related vs. Non-STEM models when applied to the test data sets. The STEM vs. STEM-Related and STEM vs. Other Degree models were very similar with acceptable predictive accuracy. The STEM-Related vs. NON-STEM model was similar to the 45^{o} diagonal line demonstrating the model had no real ability to discriminate between these outcomes and comparable predictive accuracy could be achieved by flipping a fair coin. The ROC curve for the STEM & STEM-Related vs. Non-STEM model lies in the middle of the other curves showing it had poor discrimination performance. These results indicate that greater predictive accuracy is achieved by keeping STEM as a narrowly defined category rather than expanding it to include the majors comprising the STEM-Related category. However, it also demonstrates that the students that may be the most promising targets for encouragement towards a STEM major are those already on a path towards studying a STEM-Related topic or earning a four year degree in general. These students have many of the same characteristics as the STEM students. The hierarchical element of the modeling shows how different the groups are from STEM and reflects what one might assume without this study. While some students on a Non-Degree trajectory might be persuaded to continue their education and pursue a STEM degree, the potential for converting such students is likely to be lower than that for the STEM-Related or Non-STEM students.

4. Discussion

The classification of the records by educational outcome permitted many two-outcome pairs including combinations of outcomes to be modeled. The stability of the models fitted by the multiple random samples was very good. Logistic regression models generally have greater predictive ability when the two outcomes are sharply different from one another. The predictive accuracy of the logistic regression models created in this research varied from negligible for the STEM-Related vs. Non-STEM model to outstanding for the STEM vs. Sub-4 Year Degree and STEM vs. No Degree models. The results for this application indicated a hierarchical relationship between the outcomes with STEM students exhibiting significant differences from the other students. The patterns of predictive accuracy for the models suggested that the five basic educational outcomes outlined in this research may be considered an ordered set as follows:

{STEM Degree, STEM-Related Degree, Non-STEM Degree, Sub-4 Year Degree, No Degree }

The last two categories are included for completeness in studying potential educational outcomes. This permits other researchers to compare their results using different combinations of outcomes to suit their particular focus. Considering this set to have an ordinal scale, the modeling accuracy improved as the two potential outcomes differed. For example, the models which predicted the probability of a STEM outcome improved as the alternative outcome modeled changed from a closely related category like STEM-Related to a more divergent category like Sub-4 Year Degree. The STEM-Related vs. No-Degree model had excellent predictive accuracy while the STEM-Related vs. Non-STEM model had little more predictive accuracy than tossing a fair coin. The discrimination between STEM & STEM-Related and Non-STEM in this model was little better than tossing a coin was and therefore considered "poor."

A logical extension of this analysis is to continue to perform a detailed examination of what majors should be included in STEM. New fields of study have arisen since the NELS:88 study was concluded. Disciplines such as bioengineering and nano-technology were not included in the potential majors for the NELS:88 study. Since they were not reflected in the NELS:88 study they were not tested within this analysis. However, they and others that have arisen since the NELS:88 study was concluded should be considered for STEM inclusion. We would suggest that these two fields and potentially others should be considered STEM. Other majors that have existed for some time were also not included in the set of potential majors for the NELS:88 study including systems engineering, operations research, and biophysics. Extending the classification to cover additional fields of study would aid its applicability for future research.

Another aspect of the classification that could be refined further is the STEM-Related category. This category was chosen to be a representation of majors in the "gray area" between STEM and Non-STEM and to determine if these outcomes should more properly be considered STEM, Non-STEM, or a separate group. The findings suggested that the students studying these majors are more similar to the Non-STEM students than to the STEM students. However, there are a numerous ways to decide what majors to include in the STEM-Related category. It is possible to define STEM-Related differently to move it closer to or further away from STEM and Non-STEM along the proposed ordinal scale. For example, more of the "gray area" majors could be placed in the Non-STEM category and smaller set of STEM-Related majors would remain to be used in repeating the modeling. Conversely, several of the "gray area" majors could be placed in the STEM category and the remaining set of STEM-Related majors could be examined by repeating the modeling. An additional extension of this analysis would be aggregating the "gray area" majors into sub-classes and taking out one sub-class at a time before repeating the modeling tests. For example, if the "gray area" majors were aggregated into five sub-classes a total of 5 factorial combinations could be examined. Each of these potential extensions would provide additional information for deciding how to classify majors in future studies. Consistent and clearly outlined classification schemes will better enable researchers to make more accurate comparisons.

5. Conclusions

Prior research has used classification schemes that tended to include majors inconsistently or were limited by the outcomes represented in the available data. For example, if the data only included students earning four year college majors then other outcomes were not considered. Some of the prior research involved impressively detailed and logical classification schemes for college majors, but different research questions were under study so the testing of the schemes for other applications was not considered. This study offers a methodology for researchers to identify and select the specific educational outcomes of interest and compare analytical results to those of prior studies.

The findings in this paper indicate that the method for evaluating the proposed classification scheme is a valuable tool for engineering educational research since it allows a researcher to test any classification scheme used in prior research or developed independently. Other research questions may be better examined using a classification scheme with a different design. Subsequent analysis is not limited to the proposed classification scheme utilized for this research. The methodology for modeling outcome pairs and evaluating the predictive accuracy to discern which outcomes are more closely related can be applied to any researcher's outcome classification scheme.

The comprehensive nature of this proposed classification scheme for educational outcomes is such that it allows future researchers to logically include or exclude particular groupings of majors and outcomes. The precision and flexibility permit a researcher to directly compare new models to that of prior researchers as long as the outcomes that were included in the prior research can be clearly determined. The logic for classifying new student records can be applied to new datasets and tested in different applications. This logical and consistent approach should aid future research.

There are a variety of potential applications of this research. Recruitment and retention in STEM majors are of significant public policy concern. A precise definition of STEM could aid in measuring the efficiency and effectiveness of different strategies to increase pro-STEM recruiting and retention strategies. What may work well with engineers may not work as well with scientists. Modeling the entrance of students into the various STEM majors, their academic progression through STEM, their migration between programs, their time to graduation, and the eventual number graduating all depend upon the accuracy of the STEM definition. The STEM pipeline can be improved through recruiting capable students, offering them supportive interventions if they encounter difficulties, and assisting them in transfers to other STEM majors to improve overall retention of potential STEM graduates.

Acknowledgements—This research was supported in part by a grant from the National Science Foundation (Math and Science Partnership EHR-0227016) entitled "System-wide Change for All Learners and Educators" (SCALE).

The authors also wish to gratefully acknowledge the valuable support received from the National Center of Educational Statistics (NCES) in making available access to the NELS:88 dataset. In addition, several NCES staff members including Dr. Cliff Adelman and Mr. Jeffrey Owings were very helpful in assisting with questions about the dataset.

References

1. D. Grasso, K. M. Callahan and S. Doucett, Defining Engineering Thought, *International Journal of Engineering Education*, **20**(3), 2004, pp. 412–415.
2. F. G. Splitt, Engineering Education Reform: Signs of Progress, *International Journal of Engineering Education*, **20**(6), 2004, pp. 1005–1011.
3. L. C. Benson, K. Becker, M. M. Cooper, O. H. Griffin and K. A. Smith, Engineering Education: Departments, Degrees, and Directions, *International Journal of Engineering Education*, **26**(5), 2010, pp. 1042–1048.
4. Committee on the Engineer of 2020, Phase II. *Educating the Engineer of 2020: Adapting Engineering Education to the New Century*. Committee on Engineering Education, National Academy of Engineering, The National Academies Press, 2005. Retrieved from http://www.nap.edu/openbook.php?isbn=0309096499. Accessed May 16, 2013.
5. V. Wadhwa, G. Gereffi, B. Rissing and R. Ong, Where the Engineers Are, *Issues in Science and Technology*, **23**(3), Spring, 2007, pp. 73–84.
6. G. Gereffi, V. Wadhwa, B. Rissing and R. Ong, Getting the Numbers Right: International Engineering Education in the United States, China, and India, *Journal of Engineering Education*, **97**(1), January 2008, pp. 13–25.
7. President's Council of Advisors on Science and Technology (PCAST), S. Olson, D. G. Riordan, Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics, Executive Office of the President, 2012 130 pp.
8. C. Adelman, Women and Men of the Engineering Path: A model for Analyses of Undergraduate Careers, PLLI 98-8055, U.S. Department of Education, Office of Educational Research and Improvement, Washington, DC: Government Printing Office, 1998.
9. E. Seymour and N. M. Hewitt, *Talking About Leaving: Why Undergraduates Leave the Sciences*, Westview Press, Boulder, Colorado, 1997.
10. F. L. Smyth, Ethnic and Gender Differences in Science Graduation at Selective Colleges with Implications for Admission Policy and College Choice, 2000, from an unpublished thesis for Master of Arts in Psychology, University of Virginia.
11. F. L. Smyth and J. J. McArdle, Ethnic and Gender Differences in Science Graduation at Selective Colleges with Implications for Admission Policy and College Choice, *Research in Higher Education*, **45**(4), June 2004, pp. 353–381.
12. College and Beyond Database, Andrew W. Mellon Foundation, http://www.mellon.org/news_publications/annual-reports-essays/presidents-reports/content1997#below. Accessed May 15, 2013.
13. Andrew W. Mellon Foundation, founded 1969, New York City, NY, http://www.mellon.org/about_foundation/history. Accessed May 15, 2013.
14. W. G. Bowen and Derek Bok, *The Shape of the River: Long-term consequences of considering race in college and university admissions*, Princeton University Press, Princeton, NJ, 1998.
15. Higher Education Research Institute (HERI) Cooperative Institutional Research Program (CIRP), 1966. http://www.heri.ucla.edu/cirpoverview.php, Accessed May 23, 2013.
16. A. W. Astin and H. S. Astin, *Undergraduate Science Education: The impact of different college environments on the educational pipeline in the sciences: final report*, Higher Education Research Institute, Graduate School of Education, University of California, Los Angeles, CA, 1992, 384 pp. #ES362404.
17. T. L. Hilton, J. Hsia, D .G. Solorzano and N.L. Benton, Persistence in science of high-ability minority students, Princeton, NJ: Educational Testing Service, 1989.
18. State Educational Technology Directors Association (SETDA), Science, Technology, Engineering, & Math, September 2008, pp 1–17.
19. N. Tsupros, R. Kohler, and J. Hallinen, STEM Education: A project to identify the missing components, Intermediate 1: Center for STEM Education and Leonard Gelfand Center for Service Learning and Outreach, Carnegie Mellon University, PA., 2009.
20. H. B. Lantz, Jr., Science, Technology, Engineering, and Mathematics (STEM) Education. What Form? What Function? What is STEM Education?, 2009, Retrieved from http://www.currtechintegrations.com/pdf/STEMEducationArticle.pdf , Accessed May 16, 2013.
21. R. Brown, J. Brown, K. Reardon and C. Merrill, Understanding STEM: Current Perceptions, *Technology and Engineering Teacher*, **70**(6), 2011, pp. 5–9.
22. C. Merrill, The future of TE masters degrees: STEM, Presented at the 70th Annual International Technology Education Association Conference, Louisville, Kentucky, 2009.
23. T. Kelley, Staking the Claim for the 'T' in STEM, *Journal of Technology Studies*, **36**(1), Spring, 2010, pp. 2–11.
24. L. Barakos, V. Lujan, C. Strang (2012), Science, Technology,

Engineering, Mathematics (STEM): Catalyzing change amid the confusion. Portsmouth, NH: RMC Research Corporation, Center on Instruction. Retrieved from http://www.centeroninstruction.org/files/STEM%20-%20 Catalyzing%20-Change%20Amid%20the%20Confusion.pdf. Accessed May 16, 2013.
25. J. Gerlach, STEM: Defying a simple definition, NSTA Reports, April 11, 2012, p. 3. Arlington, VA: National Science Teachers Association. Retrieved from http://www.nsta.org/publications/news/story.aspx?id=59305. Accessed May 16, 2013.
26. T. Lewis, Engineering Education in Schools, *International Journal of Engineering Education*, **23**(5), pp. 843–852, 2007.
27. National Research Council (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press. Retrieved from http://www.nap.edu/catalog.php?record_id=13165. Accessed May 16, 2013.
28. D. A. Koonce, J. Zhou, C. D. Anderson, D. A. Hening and V. Martin Conley, What is STEM?, *Proceedings* of the American Society of Engineering Education (ASEE) Annual Conference, Vancouver, BC, Canada, June 2011, pp. 1–13.
29. E. C. Kokkelenberg and E. Sinha, Who succeeds in STEM studies? An analysis of Binghamton University undergraduate students, Economics of Education Review, **29**(6), 2010, pp. 935–946.
30. G. M. Nicholls, H. Wolfe, M. E. Besterfield-Sacre and L. J. Shuman, Predicting STEM Degree Outcomes Based on Eighth Grade Data and Standard Test Scores, *Journal of Engineering Education*, **99**(3), July 2010, pp. 209–223.
31. National Science Foundation, National Center for Science and Engineering Statistics. 2011. *Science and Engineering Degrees: 1966–2008.* Detailed Statistical Tables NSF 11-316. Arlington, VA. Available at http://www.nsf.gov/statistics/nsf11316/. Accessed May 15, 2013.
32. National Center for Education Statistics, U.S. Department of Education, Washington, D.C., http://nces.ed.gov/. Accessed May 15, 2013.
33. National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), Project Officers: Elise Christopher and Jeffrey T. Owings (Washington, DC), http://nces.ed.gov/surveys/nels88/index.asp_. Accessed May 15, 2013.
34. National Center for Education Statistics, Education Longitudinal Study of 2002/06 (ELS:02/06), Project Officers: Elise Christopher, Isaiah O'Rear, and Jeffrey T. Owings (Washington, DC). http://nces.ed.gov/surveys/els2002/. Accessed May 15, 2013.
35. X. Chen and T. Weko, Students Who Study Science, Technology, Engineering, and Mathematics (STEM) in Postsecondary Education, *Statistics in Brief*, July 2009, pp. 1–24, National Center for Education Statistics, Project Officer Aurora D'Amico (Washington, DC), http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2009161. Accessed May 15, 2013.
36. National Center for Education Statistics, U.S. Department of Education, Washington, D.C., Web Tables STEM in Postsecondary Education: Entrance, Attrition, and Coursetaking Among 2003–04 Beginning Postsecondary Students, October 2012, NCES 2013–152. Authors Xianglei Chen and Phoebe Ho of MPR Associates, Inc. Project Officer: Matthew Soldner. Retrieved from http://nces.ed.gov/pubs2013/2013152.pdf, Accessed May 15, 2013.
37. G. M. Nicholls, H. Wolfe, M. E. Besterfield-Sacre, L. J. Shuman, and S. Larpkiattaworn, A Method for Identifying Variables for Predicting STEM Enrollment, *Journal of Engineering Education*, **96**(1), January 2007, pp. 33–44.
38. M. Gonen, *Analyzing Receiver Operating Characteristic Curves with SAS*, 2007, SAS Publishing, Cary, North Carolina.
39. W. J. Krzanowski and D. J. Hand, *ROC Curves for Continuous Data (Chapman & Hall/CRC Monographs on Statistics & Applied Probability)*, 2009, Chapman and Hall, Boca Raton, Florida.
40. T. Fawcett, An Introduction to ROC Analysis, *Pattern Recognition Letters*, **27**(8), 2006, pp. 861–874.
41. Wikipedia contributors, Receiver operating characteristic, *Wikipedia, The Free Encyclopedia,* permanent link http://en.wikipedia.org/w/index.php?title=Receiver_operating_characteristic&oldid=579576845 (accessed November 3, 2013).
42. D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression 2nd edition*, John Wiley & Sons, Inc, New York, 2000, pp. 160–164.

Gillian Nicholls is an Assistant Professor of Industrial & Systems Engineering & Engineering Management and a 2009-2010 Gray Faculty Fellow at the University of Alabama in Huntsville. Her research interests are in applying statistical analysis and optimization to engineering education, engineering economic analysis, and supply chain management. She holds the B.S. in Industrial Engineering (Lehigh University), Masters in Business Administration (Penn State University), M.S. in Industrial Engineering (University of Pittsburgh.), and Ph.D. in Industrial Engineering (University of Pittsburgh).

Harvey Wolfe is an Emeritus Professor of Industrial Engineering at the University of Pittsburgh. His research interests are in applying operations research methods to the health services, the development of models for assessing engineering education, and engineering ethics. He is the author of numerous journal articles and three books. He is a co-author of *Engineering Ethics: Balancing Cost Schedule and Risk - Lessons Learned from the Space Shuttle* (Cambridge University Press, 1997). He holds the B.E.S. in Industrial Engineering, M.S.E. in Operations Research, and Ph.D. in Operations Research, all from the Johns Hopkins University.

Mary E. Besterfield-Sacre, is an Associate Professor and Fulton C. Noss Faculty Fellow in the Department of Industrial Engineering and the Director for the Engineering Education Research Center (EERC) in the Swanson School of Engineering at the University of Pittsburgh. Her principal research is in engineering education assessment, which has been funded by the NSF, Department of Education, Sloan Foundation, Engineering Information Foundation, and NCIIA. Dr. Sacre's current research focuses on three distinct but highly correlated areas—innovative design and entrepreneurship, engineering modeling, and global competency in engineering. She has served as an associate editor for the *Journal of Engineering Education* and is currently associate editor for the *Applications in Engineering Education Journal*. She received her B.S. in Engineering Management from the University of Missouri—Rolla, her M.S. in Industrial Engineering from Purdue University, and a Ph.D. in Industrial Engineering at the University of Pittsburgh.

Larry J. Shuman is Senior Associate Dean for Academic Affairs and Distinguished Service Professor of Industrial Engineering at the University of Pittsburgh. His research focuses on improving the engineering educational experience with an emphasis on assessment of design and problem solving, and the study of the ethical behavior of engineers and engineering managers. A former senior editor of the *Journal of Engineering Education*, Dr. Shuman is the founding editor of *Advances in Engineering Education.* He has published widely in the engineering education literature, and is co-author of *Engineering Ethics: Balancing Cost, Schedule and Risk—Lessons Learned from the Space Shuttle* (Cambridge University Press, 1997). He received his Ph.D. from The Johns Hopkins University in Operations Research and the BSEE from the University of Cincinnati. He is an ASEE Fellow.